FORM PTO-1390 (REV. 12-2001) U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE

TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371

ATTORNEY'S DOCKET NUMBER: MWH-0006US

U.S. APPLICATION NO. (If known, see 37 CFR 1.5): 10/049407

INTERNATIONAL APPLICATION NO.	INTERNATIONAL FILING DATE	PRIORITY DATE CLAIMED
PCT/US00/40519	01 August 2000 (1.08.00)	06 August 1999 (6.08.99)

TITLE OF INVENTION
Drug Target Isogenes: Polymorphisms in the 5-Hydroxytryptamine Receptor 1A Gene

APPLICANT(S) FOR DO/EO/US DENTON, R. Rex; KLIEM, Stefanie E.; NANDABALAN, Krishnan and STEPHENS, J. Claiborne

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. [X] This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. [ ] This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. [X] This is an express request to begin national examination procedures (35 U.S.C. 371(f)). The submission must include items (5), (6), (9) and (21) indicated below.
4. [ ] The US has been elected by the expiration of 19 months from the priority date (Article 31).
5. [X] A copy of the International Application as filed (35 U.S.C. 371(c)(2))
   a. [ ] is attached hereto (required only if not communicated by the International Bureau).
   b. [ ] has been communicated by the International Bureau.
   c. [X] is not required, as the application was filed in the United States Receiving Office (RO/US).
6. [ ] An English language translation of the International Application as filed (35 U.S.C. 371(c)(2)).
   a. [ ] is attached hereto.
   b. [ ] has been previously submitted under 35 U.S.C. 154(d)(4).
7. [X] Amendments to the claims of the International Aplication under PCT Article 19 (35 U.S.C. 371(c)(3))
   a. [ ] are attached hereto (required only if not communicated by the International Bureau).
   b. [ ] have been communicated by the International Bureau.
   c. [ ] have not been made; however, the time limit for making such amendments has NOT expired.
   d. [X] have not been made and will not be made.
8. [ ] An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371 (c)(3)).
9. [X] An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
10. [ ] An English lanugage translation of the annexes of the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11 to 20 below concern document(s) or information included:

11. [X] An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. [ ] An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
13. [ ] A FIRST preliminary amendment.
14. [ ] A SECOND or SUBSEQUENT preliminary amendment.
15. [ ] A substitute specification.
16. [ ] A change of power of attorney and/or address letter.
17. [X] A computer-readable form of the sequence listing in accordance with PCT Rule 13ter.2 and 35 U.S.C. 1.821 - 1.825.
18. [ ] A second copy of the published international application under 35 U.S.C. 154(d)(4).
19. [ ] A second copy of the English language translation of the international application under 35 U.S.C. 154(d)(4).
20. [X] Other items or information:

(a) Request for Expedited Examination under 37 C.F.R. 1.496(b)
(b) The claims in the International Application were modified by a PCT Article 34 Amendment, a copy of which is not required, as the amendment was filed in the IPEA/US. It is Applicants intent that the national stage processing be performed on the same claims in the International Application as amended by the Article 34 Amendment.

JC11 Rec'd PCT/PTO 06 FEB 2002

U.S. APPLICATION NO. (if known, see 37 CFR 1.5)	INTERNATIONAL APPLICATION NO.	ATTORNEY'S DOCKET NUMBER
10/049107	PCT/US00/40519	MWH-0006US

21 [X] The following fees are submitted:	CALCULATIONS PTO USE ONLY	
BASIC NATIONAL FEE (37 CFR 1.492 (a) (1)-(5)):		
Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO **$1040.00**		
International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO **$890.00**		
International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO **$740.00**		
International preliminary examination fee (37 CFR 1.482) paid to USPTO but all claims did not satisfy provisions of PCT Article 33(1)-(4) **$710.00**		
International preliminary examination fee (37 CFR 1.482) paid to USPTO and all claims satisfied provisions of PCT Article 33(1)-(4) **$100.00**		
ENTER APPROPRIATE BASIC FEE AMOUNT =	$ 100.00	
Surcharge of **$130.00** for furnishing the oath or declaration later than [ ] 20 [ ] 30 months from the earliest claimed priority date (37 CFR 1.492(e)).	$ 0.00	

CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE		
Total claims	25 - 20 =	5	x **$18.00**	$ 90.00	
Independent claims	11 - 3 =	8	x **$84.00**	$ 672.00	
MULTIPLE DEPENDENT CLAIM(S) (if applicable)			+ **$280.00**	$ 0.00	
TOTAL OF ABOVE CALCULATIONS =				$ 862.00	
[ ] Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by 1/2. +				$ 0.00	
SUBTOTAL =				$ 862.00	
Processing fee of **$130.00** for furnishing the English translation later than [ ] 20 [ ] 30 months from the earliest claimed priority date (37 CFR 1.492(f)).				$ 0.00	
TOTAL NATIONAL FEE =				$ 862.00	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). **$40.00** per property +				$ 0.00	
TOTAL FEES ENCLOSED =				$ 862.00	
				Amount to be refunded:	$
				charged:	$

a. [ ] A check in the amount of $ _____ to cover the above fees is enclosed.

b. [X] Please charge my Deposit Account No. 50-1293 in the amount of $ 862.00 to cover the above fees. A duplicate copy of this sheet is enclosed.

c. [X] The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 50-1293 A duplicate copy of this sheet is enclosed.

d. [X] Fees are to be charged to a credit card. **WARNING:** Information on this form may become public. **Credit card information should not be included on this form.** Provide credit card information and authorization on PTO-2038.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137 (a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:
Gisela M. Field
Genaissance Pharmaceuticals
5 Science Park
New Haven, CT 06511

Customer Number: 25106

Gisela M. Field
SIGNATURE

Gisela M. Field
NAME

47,562
REGISTRATION NUMBER

WO 01/10884 PCT/US00/40519

Rec'd PCT/PTO 06 FEB 2002

DRUG TARGET ISOGENES:

POLYMORPHISMS IN THE 5-HYDROXYTRYPTAMINE RECEPTOR 1A GENE

RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Application Serial No. 60/147,711 filed August 06, 1999.

FIELD OF THE INVENTION

This invention relates to variation in genes that encode pharmaceutically important proteins. In particular, this invention provides genetic variants of the human 5-hydroxytryptamine receptor 1A gene (HTR1A) gene and methods for identifying which variant(s) of this gene is/are possessed by an individual.

BACKGROUND OF THE INVENTION

Current methods for identifying pharmaceuticals to treat disease often start by identifying, cloning, and expressing an important target protein related to the disease. A determination of whether an agonist or antagonist is needed to produce an effect that may benefit a patient with the disease is then made. Then, vast numbers of compounds are screened against the target protein to find new potential drugs. The desired outcome of this process is a drug that is specific for the target, thereby reducing the incidence of the undesired side effects usually caused by a compound's activity at non-intended targets.

What this approach fails to consider, however, is that natural variability exists in any and every population with respect to a particular protein. A target protein currently used to screen drugs typically is expressed by a gene cloned from an individual who was arbitrarily selected. However, the nucleotide sequence of a particular gene may vary tremendously among individuals. Subtle alteration(s) in the primary nucleotide sequence of a gene encoding a target protein may be manifested as significant variation in expression of or in the structure and/or function of the protein. Such alterations may explain the relatively high degree of uncertainty inherent in treatment of individuals with drugs whose design is based upon a single representative example of the target. For example, it is well-established that some classes of drugs frequently have lower efficacy in some individuals than others, which means such individuals and their physicians must weigh the possible benefit of a larger dosage against a greater risk of side effects. In addition, variable information on the biological function or effects of a particular protein may be due to different scientists unknowingly studying different isoforms of the gene encoding the protein. Thus, information on the type and frequency of genomic variation that exists for pharmaceutically important proteins would be useful.

The organization of single nucleotide variations (polymorphisms) in the primary sequence of a gene into one of the limited number of combinations that exist as units of inheritance is termed a

haplotype. Each haplotype therefore contains significantly more information than individual unorganized polymorphisms. Haplotypes provide an accurate measurement of the genomic variation in the two chromosomes of an individual.

It is well-established that many diseases are associated with specific variations in gene sequences. However while there are examples in which individual polymorphisms act as genetic markers for a particular phenotype, in other cases an individual polymorphism may be found in a variety of genomic backgrounds and therefore shows no definitive coupling between the polymorphism and the causative site for the phenotype (Clark AG et al. 1998 *Am J Hum Genet* **63**:595-612; Ulbrecht M et al. 2000 *Am J Respir Crit Care Med* **161**: 469-74). In addition, the marker may be predictive in some populations, but not in other populations (Clark AG et al. 1998 *supra*). In these instances, a haplotype will provide a superior genetic marker for the phenotype (Clark AG et al. 1998 *supra*; Ulbrecht M et al. 2000, *supra*; Ruaño G & Stephens JC *Gen Eng News* **19** (21), December 1999).

Analysis of the association between each observed haplotype and a particular phenotype permits ranking of each haplotype by its statistical power of prediction for the phenotype. Haplotypes found to be strongly associated with the phenotype can then have that positive association confirmed by alternative methods to minimize false positives. For a gene suspected to be associated with a particular phenotype, if no observed haplotypes for that gene show association with the phenotype of interest, then it may be inferred that variation in the gene has little, if any, involvement with that phenotype (Ruaño & Stephens 1999, *supra*). Thus, information on the observed haplotypes and their frequency of occurrence in various population groups will be useful in a variety of research and clinical applications.

One possible drug target for the treatment of neuropsychiatric diseases and Tourette's syndrome is the 5-hydroxytryptamine receptor 1A gene (HTR1A) gene or its encoded product. The 5-hydroxytryptamine receptor 1A (HTR1A) is one of fourteen known receptors for 5-hydroxytryptamine (5-HT; serotonin), a biogenic hormone that functions as a neurotransmitter, a hormone, and a mitogen. Serotonergic pathways are believed to participate in a wide variety of physiological and behavioral processes, including nociception, thermoregulation, cardiovascular function, feeding, aggression, and stress response. Aberrant function in the serotonergic pathway has been implicated in neuropsychiatric diseases such as anxiety, depression, substance abuse, and Tourette's syndrome. Genetic predisposition to these disorders may involve variation in genes encoding 5-HT receptors.

Serotonin receptors belong to the superfamily of seven-transmembrane domain, G-protein-coupled receptors. The 14 members of the 5-HT receptor family have been classified extensively on the basis of molecular biological, pharmacological, biochemical, and physiological properties. The most thoroughly studied to date is HTR1A. The activity of HTR1A is mediated by G proteins that inhibit the activity of adenylate cyclase.

5HT1A receptors are expressed in a tissue specific manner, predominantly in the dorsal and median raphe' hippocampus, septal nuclei, amygdala, and certain cortical layers. Both presynaptic and

postsynaptic expresssion on serotonergic neurons is found. Presynaptic HTR1A is a classical "autoreceptor" in the raphe' nuclei. Presynaptic serotonergic neurons synthesize the majority of brain serotonin, and activation of HTR1A molecules on these neurons completes a negative feedback loop, inhibiting serotonin synthesis, turnover, and release, thus attenuating system-wide serotonergic signaling. This important regulatory action provides a target for clinical intervention, and has been successfully exploited with HTR1A partial agonists, such as the anxiolytics, buspirone and gepirone. In addition, antagonists of HTR1A are thought to enhance the antidepressant action of selective serotonin reuptake inhibitors (SSRIs), such a fluoxetine. SSRIs function by blocking 5-HT binding to the serotonin transporter in the synaptic cleft, thereby elevating the levels of synaptic serotonin. Antagonist blockade of the inhibitory presynaptic HTR1A augments and accelerates the effectiveness of SSRIs. Postsynaptic HTR1A is found primarily in the limbic region (hippocampus, septum) and cerebral cortex, regions where the serotonergic neurons project. Activation of postsynaptic HTR1A is thought to induce hyperpolarization of the postsynaptic neuron via activation of a G-protein coupled potassium channel. The result is decreased neurotransmitter release and decreased neuron firing rate.

The 5-hydroxytryptamine receptor 1A gene gene is located on chromosome 5q11.2-q13 (Kobilka et al., *Nature*. 1987 329:75-79) and contains 1 exon that encodes a 422 amino acid protein. Although a full genomic sequence has not been published, a reference sequence for the HTR1A gene can be assembled from noncontiguous published sequences. Overlapping sequences for the promoter region and the 5'end of the exon have been published and are shown in Figure 1 (GenBank Accession No. Z11168.1; SEQ ID NO:1), which includes the promoter and 5'end of the exon, and Figure 2 (GenBank AccessionNo. M83181.1; SEQ ID NO:2), which includes the complete exon. A reference sequence for the HTR1A gene is shown in Figure 3 (SEQ ID NO:3; nucleotides 1-784 of GenBank Accession No. Z11168.1 and nucleotides 1-1938 of GenBank Accession No. M83181.1). Reference sequences for the coding sequence and protein are shown in Figures 4 (SEQ ID NO:4) and 5 (SEQ ID NO:5), respectively.

Significant features of the HTR1A gene include: a lack of typical TATA box elements in the promoter, but instead, extensive runs of GC-rich sequence, particularly the consensus GGGGC/AGGGG sequence, which is indicative of MAZ and Sp1 transcription factor binding sites (Parks and Shenk, *J. Biol. Chem.* 271: 4417-4430, 1996). Significant features of the HTR1A protein sequence (Fig. 4) include an amino terminus extracellular domain at amino acids (AA) 1-36; seven transmembrane domains at AA 37-62, AA 74-98, AA 110-132, AA153-178, AA 192-217, AA 346-367, and AA 379-403; and a carboxyl terminus cytoplasmic domain at AA 404-422.

Polymorphisms at the HTR1A locus, either identified as restriction fragment length polymorphisms (RFLPs), including TaqI, RsaI, and SacI RFLPs, or as single nucleotide polymorphisms (SNPs), have been previously reported. Melmer et al. (*Genomics*. 1991 11:767-769) identified a RFLP for the enzyme TaqI which was used to refine the map location of HTR1A and demonstrated tight linkage with several highly polymorphic markers in reference populations assembled for association studies of

schizophrenia. In association studies between schizophrenia and the serotonin receptors HTR1A and HTR2A, positive association was found with HTR2A on the long arm of chromosome 13, but not with HTR1A (Inayama et al., *Am. J. Med. Genet.* 1996 67:103-105). In this study, a RsaI RFLP for HTR1A was used, originally identified by Warren et al., *Hum. Mol. Genet.* 1: 778, 1992. The RsaI RFLP corresponds to a silent polymorphism, a guanine to adenine change at nucleotide position 685 (amino acid 98; Fig 4) identified by Xie et al. (*Neuropsychopharmacology* 12: 263-268. 1995). This polymorphism has also been reported in the NCBI SNP Database; (Ref SNP #6354) as of July 24, 2000. A third RFLP for the enzyme SacI is also associated with the HTR1A locus (Khan et al., *Nucl. Acids Res.* 18:691, 1990), but the polymorphic site is probably found in the flanking regions of the gene, since no SacI sites are found in the reference nucleotide sequence. Additionally, a single nucleotide polymorphism corresponding to a polymorphism of cytosine or thymine at nucleotide position 943 in Figure 2 has been reported in the Human Genic Bi-Allelic Sequences Database (HGBase:SNP000002698; Kawanishi et al., *Am. J. Med. Genet.* 1998. 81:434-439) as of July 28, 2000.

There have been several reports of SNPs that produce variation in the amino acid sequence. The following SNPs in Figure 2 cause amino acid variations in the HTR1A protein: a polymorphism of cytosine or thymine at nucleotide position 438 which results in proline or leucine at AA position 16 (Kawanishi et al., 1998. Am. J. Med. Genet. 81:434-9), a polymorphism of guanine or adenine at nucleotide position 455 which results in glycine or serine at AA position 22 (Nakhai et al., *Biochem Biophys. Res. Commun.* 210: 530-536 1995), a polymorphism of adenine or guanine at nucleotide position 473 leading to isoleucine or valine at AA position 28 (Nakhai et al., *supra*), guanine or thymine at nucleotide position 1050, which produces arginine or leucine at AA 220, (Lam et al., *Biochem. Biophys. Res. Commun.* 219: 853-858, 1996), a polymorphism of cytosine or guanine at nucleotide position 1645 which produces lysine or asparagine at AA 420 (Lam et al., *supra*), and a polymorphism of guanine or adenine at nucleotide position 1209 which results in glycine or aspartate at AA 273 (HGBase :SNP0000002697; July 28, 2000)

Because of the potential for polymorphisms in the HTR1A gene to affect the expression and function of the encoded protein, it would be useful to determine whether additional polymorphisms exist in the HTR1A gene, as well as how such polymorphisms are combined in different copies of the gene. Such information would be useful for studying the biological function of the HTR1A receptor as well as in identifying agonists and/or antagonists targeting this receptor in the treatment of behavioral disorders, such as anxiety and depression.

SUMMARY OF THE INVENTION

Accordingly, the inventors herein have discovered 3 novel polymorphic sites in the HTR1A gene. These polymorphic sites (PS) correspond to the following nucleotide positions in Figure 3: 996 (PS1), 1598 (PS5) and 1639 (PS6). The polymorphisms at these sites are cytosine or thymine at PS1, cytosine or

thymine at PS5 and cytosine or guanine at PS6. In addition, the inventors have determined the identity of the alternative nucleotides present at these sites, as well as at the previously identified sites at nucleotides 1222 (PS2), 1257 (PS3), 1469 (PS4), 1727 (PS7), 1834 (PS8) and 1993 (PS9), in a human reference population of 79 unrelated individuals self-identified as belonging to one of four major population groups: African descent, Asian, Caucasian and Hispanic/Latino. It is believed that HTR1A-encoding polynucleotides containing one or more of the novel polymorphic sites reported herein will be useful in studying the expression and biological function of HTR1A, as well as in developing drugs targeting this protein. In addition, information on the combinations of polymorphisms in the HTR1A gene may have diagnostic and forensic applications.

Thus, in one embodiment, the invention provides an isolated polynucleotide comprising a nucleotide sequence which is a polymorphic variant of a reference sequence for the HTR1A gene or a fragment thereof. The reference sequence comprises SEQ ID NO:3 and the polymorphic variant comprises at least one polymorphism selected from the group consisting of thymine at PS1, thymine at PS5, and guanine at PS6. In a preferred embodiment, the polymorphic variant comprises one or more additional polymorphisms selected from the group consisting of thymine at PS2, guanine at PS3, adenine at PS4, thymine at PS7, thymine at PS8 and adenine at PS9. A particularly preferred polymorphic variant is a naturally-occurring isoform (also referred to herein as an "isogene") of the HTR1A gene. An HTR1A isogene of the invention comprises cytosine or thymine at PS1, cytosine or thymine at PS2, adenine or guanine at PS3, guanine or adenine at PS4, cytosine or thymine at PS5, cytosine or guanine at PS6, cytosine or thymine at PS7, guanine or thymine at PS8 and guanine or adenine at PS9 . The invention also provides a collection of HTR1A isogenes, referred to herein as an HTR1A genome anthology.

An HTR1A isogene may be defined by the combination and order of these polymorphisms in the isogene, which is referred to herein as an HTR1A haplotype. Thus, the invention also provides data on the number of different HTR1A haplotypes found in the above four population groups. This haplotype data is useful in methods for deriving an HTR1A haplotype from an individual's genotype for the HTR1A gene and for determining an association between an HTR1A haplotype and a particular trait.

In another embodiment, the invention provides a polynucleotide comprising a polymorphic variant of a reference sequence for an HTR1A cDNA or a fragment thereof. The reference sequence comprises SEQ ID NO:4 (Fig. 4) and the polymorphic cDNA comprises at least one polymorphism selected from the group consisting of thymine at a position corresponding to nucleotide 423 and guanine at a position corresponding to nucleotide 464. In a preferred embodiment, the polymorphic variant comprises one or more polymorphisms selected from the group consisting of thymine at a position corresponding to nucleotide 47, guanine at a position corresponding to nucleotide 82, adenine at a position corresponding to nucleotide 294, thymine at a position corresponding to nucleotide 552, thymine at a position corresponding to nucleotide 659 and adenine at a position corresponding to nucleotide 818.

Polynucleotides complementary to these HTR1A genomic and cDNA variants are also provided

by the invention.

In other embodiments, the invention provides a recombinant expression vector comprising one of the polymorphic genomic variants operably linked to expression regulatory elements as well as a recombinant host cell transformed or transfected with the expression vector. The recombinant vector and host cell may be used to express HTR1A for protein structure analysis and drug binding studies.

In yet another embodiment, the invention provides a polypeptide comprising a polymorphic variant of a reference amino acid sequence for the HTR1A protein. The reference amino acid sequence comprises SEQ ID NO:5 (Fig. 5) and the polymorphic variant comprises the variant amino acid glycine at a position corresponding to amino acid position 155. In some embodiments, the polymorphic variant also comprises at least one variant amino acid selected from the group consisting of leucine at a position corresponding to amino acid position 16, valine at a position corresponding to amino acid position 28, leucine at a position corresponding to amino acid position 220 and aspartic acid at a position corresponding to amino acid position 273. A polymorphic variant of HTR1A is useful in studying the effect of the variation on the biological activity of HTR1A as well as studying the binding affinity of candidate drugs targeting HTR1A for the treatment of neuropsychiatric diseases and Tourette's syndrome.

The present invention also provides antibodies that recognize and bind to the above polymorphic HTR1A protein variant. Such antibodies can be utilized in a variety of diagnostic and prognostic formats and therapeutic methods.

In other embodiments, the invention provides methods, compositions, and kits for haplotyping and/or genotyping the HTR1A gene in an individual. The methods involve identifying the nucleotide or nucleotide pair present at one or more polymorphic sites selected from PS1, PS5 and PS6 in one or both copies of the HTR1A gene from the individual. The compositions contain oligonucleotide probes and primers designed to specifically hybridize to one or more target regions containing, or that are adjacent to, a polymorphic site. The methods and compositions for establishing the genotype or haplotype of an individual at the novel polymorphic sites described herein are useful for studying the effect of the polymorphisms in the etiology of diseases affected by the expression and function of the HTR1A protein, studying the efficacy of drugs targeting HTR1A, predicting individual susceptibility to diseases affected by the expression and function of the HTR1A protein and predicting individual responsiveness to drugs targeting HTR1A.

In yet another embodiment, the invention provides a method for identifying an association between a genotype or haplotype and a trait. In preferred embodiments, the trait is susceptibility to a disease, severity of a disease, the staging of a disease or response to a drug. Such methods have applicability in developing diagnostic tests and therapeutic treatments for neuropsychiatric diseases and Tourette's syndrome.

The present invention also provides transgenic animals comprising one of the HTR1A genomic polymorphic variants described herein and methods for producing such animals. The transgenic animals

are useful for studying expression of the HTR1A isogenes *in vivo*, for *in vivo* screening and testing of drugs targeted against HTR1A protein, and for testing the efficacy of therapeutic agents and compounds for neuropsychiatric diseases and Tourette's syndrome in a biological system.

The present invention also provides a computer system for storing and displaying polymorphism data determined for the HTR1A gene. The computer system comprises a computer processing unit; a display; and a database containing the polymorphism data. The polymorphism data includes the polymorphisms, the genotypes and the haplotypes identified for the HTR1A gene in a reference population. In a preferred embodiment, the computer system is capable of producing a display showing HTR1A haplotypes organized according to their evolutionary relationships.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates a partial reference sequence for the HTR1A gene (Genbank Version Number Z11168.1; contiguous lines; SEQ ID NO:1), with the start and stop positions of each region of coding sequence indicated with a bracket ([or]) and the numerical position below the sequence and the polymorphic sites and polymorphisms identified by Applicants in a reference population indicated by the variant nucleotide positioned below the polymorphic site in the seqence.

Figure 2 illustrates a partial reference sequence for the HTR1A gene (Genbank Version Number M83181.1; contiguous lines; SEQ ID NO:2), with the start and stop positions of each region of coding sequence indicated with a bracket ([or]) and the numerical position below the sequence and the polymorphic sites and polymorphisms identified by Applicants in a reference population indicated by the variant nucleotide positioned below the polymorphic site in the seqence.

Figure 3 a reference sequence for the HTR1A gene (contiguous lines; SEQ ID NO:3), which comprises nucleotides 1-784 of GenBank Accession No. Z11168.1 followed by nucleotides 1-1938 of GenBank Accession No. M83181.1, with the start and stop positions of each region of coding sequence indicated with a bracket ([or]) and the numerical position below the sequence and the polymorphic sites and polymorphisms identified by Applicants in a reference population indicated by the variant nucleotide positioned below the polymorphic site in the seqence.

Figure 4 illustrates a reference sequence for the HTR1A coding sequence (contiguous lines; SEQ ID NO:4) with the polymorphic sites and polymorphisms identified by Applicants in a reference population indicated by the variant nucleotide positioned below the polymorphic site in the seqence.

Figure 5 illustrates a reference sequence for the HTR1A protein (contiguous lines; SEQ ID NO:5) with the variant amino acids caused by the polymorphisms of Figure 4 positioned below the polymorphic site in the sequence.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention is based on the discovery of novel variants of the HTR1A gene. As

described in more detail below, the inventors herein discovered 3 novel polymorphic sites by characterizing the HTR1A gene found in genomic DNAs isolated from an Index Repository that contains immortalized cell lines from one chimpanzee and 93 human individuals. The human individuals included a reference population of 79 unrelated individuals self-identified as belonging to one of four major population groups: Caucasian (22 individuals), African descent (20 individuals) Asian (20 individuals) Hispanic/Latino (17 individuals). To the extent possible, the members of this reference population were organized into population subgroups by the self-identified ethnogeographic origin of their four grandparents as shown in Table 1 below.

Table 1. Population Groups in the Index Repository

Population Group	Population Subgroup	No. of Individuals
African descent		20
	Sierra Leone	1
Asian		20
	Burma	1
	China	3
	Japan	6
	Korea	1
	Philippines	5
	Vietnam	4
Caucasian		22
	British Isles	3
	British Isles/Central	4
	British Isles/Eastern	1
	Central/Eastern	1
	Eastern	3
	Central/Mediterranean	1
	Mediterranean	2
	Scandinavian	2
Hispanic/Latino		17
	Caribbean	7
	Caribbean (Spanish Descent)	2
	Central American (Spanish Descent)	1
	Mexican American	4
	South American (Spanish Descent)	3

In addition, the Index Repository contains three unrelated indigenous American Indians (one from each of North, Central and South America), one three-generation Caucasian family (from the CEPH Utah cohort) and one two-generation African-American family.

Using the HTR1A genotypes identified in the Index Repository and the methodology described in the Examples below, the inventors herein also determined the haplotypes found on each chromosome for most human members of this repository. The HTR1A genotypes and haplotypes found in the repository include those shown in Tables 4 and 5, respectively. The polymorphism and haplotype data disclosed herein are useful for studying population diversity, anthropological lineage, the significance of diversity

and lineage at the phenotypic level, paternity testing, forensic applications, and for identifying associations between the HTR1A genetic variation and a trait such as level of drug response or susceptibility to disease.

In the context of this disclosure, the following terms shall be defined as follows unless otherwise indicated:

Allele - A particular form of a genetic locus, distinguished from other forms by its particular nucleotide sequence.

Candidate Gene – A gene which is hypothesized to be responsible for a disease, condition, or the response to a treatment, or to be correlated with one of these.

Gene - A segment of DNA that contains all the information for the regulated biosynthesis of an RNA product, including promoters, exons, introns, and other untranslated regions that control expression.

Genotype – An unphased 5′ to 3′ sequence of nucleotide pair(s) found at one or more polymorphic sites in a locus on a pair of homologous chromosomes in an individual. As used herein, genotype includes a full-genotype and/or a sub-genotype as described below.

Full-genotype – The unphased 5′ to 3′ sequence of nucleotide pairs found at all known polymorphic sites in a locus on a pair of homologous chromosomes in a single individual.

Sub-genotype – The unphased 5′ to 3′ sequence of nucleotides seen at a subset of the known polymorphic sites in a locus on a pair of homologous chromosomes in a single individual.

Genotyping – A process for determining a genotype of an individual.

Haplotype – A 5′ to 3′ sequence of nucleotides found at one or more polymorphic sites in a locus on a single chromosome from a single individual. As used herein, haplotype includes a full-haplotype and/or a sub-haplotype as described below.

Full-haplotype – The 5′ to 3′ sequence of nucleotides found at all known polymorphic sites in a locus on a single chromosome from a single individual.

Sub-haplotype – The 5′ to 3′ sequence of nucleotides seen at a subset of the known polymorphic sites in a locus on a single chromosome from a single individual.

Haplotype pair – The two haplotypes found for a locus in a single individual.

Haplotyping – A process for determining one or more haplotypes in an individual and includes use of family pedigrees, molecular techniques and/or statistical inference.

Haplotype data - Information concerning one or more of the following for a specific gene: a listing of the haplotype pairs in each individual in a population; a listing of the different haplotypes in a population; frequency of each haplotype in that or other populations, and any known associations between one or more haplotypes and a trait.

Isoform – A particular form of a gene, mRNA, cDNA or the protein encoded thereby, distinguished from other forms by its particular sequence and/or structure.

Isogene – One of the isoforms of a gene found in a population. An isogene contains all of the

polymorphisms present in the particular isoform of the gene.

Isolated – As applied to a biological molecule such as RNA, DNA, oligonucleotide, or protein, isolated means the molecule is substantially free of other biological molecules such as nucleic acids, proteins, lipids, carbohydrates, or other material such as cellular debris and growth media. Generally, the term "isolated" is not intended to refer to a complete absence of such material or to absence of water, buffers, or salts, unless they are present in amounts that substantially interfere with the methods of the present invention.

Locus - A location on a chromosome or DNA molecule corresponding to a gene or a physical or phenotypic feature.

Naturally-occurring – A term used to designate that the object it is applied to, e.g., naturally-occurring polynucleotide or polypeptide, can be isolated from a source in nature and which has not been intentionally modified by man.

Nucleotide pair – The nucleotides found at a polymorphic site on the two copies of a chromosome from an individual.

Phased – As applied to a sequence of nucleotide pairs for two or more polymorphic sites in a locus, phased means the combination of nucleotides present at those polymorphic sites on a single copy of the locus is known.

Polymorphic site (PS) – A position within a locus at which at least two alternative sequences are found in a population, the most frequent of which has a frequency of no more than 99%.

Polymorphic variant – A gene, mRNA, cDNA, polypeptide or peptide whose nucleotide or amino acid sequence varies from a reference sequence due to the presence of a polymorphism in the gene.

Polymorphism – The sequence variation observed in an individual at a polymorphic site. Polymorphisms include nucleotide substitutions, insertions, deletions and microsatellites and may, but need not, result in detectable differences in gene expression or protein function.

Polymorphism data – Information concerning one or more of the following for a specific gene: location of polymorphic sites; sequence variation at those sites; frequency of polymorphisms in one or more populations; the different genotypes and/or haplotypes determined for the gene; frequency of one or more of these genotypes and/or haplotypes in one or more populations; any known association(s) between a trait and a genotype or a haplotype for the gene.

Polymorphism Database – A collection of polymorphism data arranged in a systematic or methodical way and capable of being individually accessed by electronic or other means.

Polynucleotide – A nucleic acid molecule comprised of single-stranded RNA or DNA or comprised of complementary, double-stranded DNA.

Population Group – A group of individuals sharing a common ethnogeographic origin.

Reference Population – A group of subjects or individuals who are predicted to be representative of the genetic variation found in the general population. Typically, the reference population represents

the genetic variation in the population at a certainty level of at least 85%, preferably at least 90%, more preferably at least 95% and even more preferably at least 99%.

Single Nucleotide Polymorphism (SNP) – Typically, the specific pair of nucleotides observed at a single polymorphic site. In rare cases, three or four nucleotides may be found.

Subject – A human individual whose genotypes or haplotypes or response to treatment or disease state are to be determined.

Treatment - A stimulus administered internally or externally to a subject.

Unphased – As applied to a sequence of nucleotide pairs for two or more polymorphic sites in a locus, unphased means the combination of nucleotides present at those polymorphic sites on a single copy of the locus is not known.

The inventors herein have discovered 3 novel polymorphic sites in the HTR1A gene. The polymorphic sites identified by the inventors are referred to as PS1-9 to designate the order in which they are located in the gene (see Table 3 below), with the novel polymorphic sites referred to as PS1, PS5 and PS6.

Thus, in one embodiment, the invention provides an isolated polynucleotide comprising a polymorphic variant of the HTR1A gene or a fragment of the gene which contains at least one of the novel polymorphic sites PS1, PS5 and PS6 described herein. The nucleotide sequence of a variant HTR1A gene is identical to the reference genomic sequence for those portions of the gene examined, as described in the Examples below, except that it comprises a different nucleotide at one or more of the novel polymorphic sites PS1, PS5 and PS6, and may also comprise one or more additional polymorphisms selected from the group consisting of thymine at PS2, guanine at PS3, adenine at PS4, thymine at PS7, thymine at PS8 and adenine at PS9. Similarly, the nucleotide sequence of a variant fragment of the HTR1A gene is identical to the corresponding portion of the reference sequence except for having a different nucleotide at one or more of the novel polymorphic sites described herein. Thus, the invention specifically does not include polynucleotides comprising a nucleotide sequence identical to the reference sequence (or other reported HTR1A sequences) or to portions of the reference sequence (or other reported HTR1A sequences), except for genotyping oligonucleotides as described below.

The location of a polymorphism in a variant gene or fragment is identified by aligning its sequence against SEQ ID NO:3. The polymorphism is selected from the group consisting of thymine at PS1, thymine at PS5 and guanine at PS6. In a preferred embodiment, the polymorphic variant comprises a naturally-occurring isogene of the HTR1A gene which is defined by any one of haplotypes 1-10 shown in Table 5 below.

Polymorphic variants of the invention may be prepared by isolating a clone containing the HTR1A gene from a human genomic library. The clone may be sequenced to determine the identity of the nucleotides at the polymorphic sites described herein. Any particular variant claimed herein could be prepared from this clone by performing *in vitro* mutagenesis using procedures well-known in the art.

HTR1A isogenes may be isolated using any method that allows separation of the two "copies" of the HTR1A gene present in an individual, which, as readily understood by the skilled artisan, may be the same allele or different alleles. Separation methods include targeted *in vivo* cloning (TIVC) in yeast as described in WO 98/01573, U.S. Patent No. 5,866,404, and copending U.S. application Serial No. 08/987,966. Another method, which is described in copending U.S. Application Serial No. 08/987,966, uses an allele specific oligonucleotide in combination with primer extension and exonuclease degradation to generate hemizygous DNA targets. Yet other methods are single molecule dilution (SMD) as described in Ruaño et al., Proc. Natl. Acad. Sci. 87:6296-6300, 1990; and allele specific PCR (Ruaño et al., 17 Nucleic Acids. Res. 8392, 1989; Ruaño et al., 19 Nucleic Acids Res. 6877-6882, 1991; Michalatos-Beloin et al., 24 Nucleic Acids Res. 4841-4843, 1996).

The invention also provides HTR1A genome anthologies, which are collections of HTR1A isogenes found in a given population. The population may be any group of at least two individuals, including but not limited to a reference population, a population group, a family population, a clinical population, and a same sex population. An HTR1A genome anthology may comprise individual HTR1A isogenes stored in separate containers such as microtest tubes, separate wells of a microtitre plate and the like. Alternatively, two or more groups of the HTR1A isogenes in the anthology may be stored in separate containers. Individual isogenes or groups of isogenes in a genome anthology may be stored in any convenient and stable form, including but not limited to in buffered solutions, as DNA precipitates, freeze-dried preparations and the like. A preferred HTR1A genome anthology of the invention comprises a set of isogenes defined by the haplotypes shown in Table 5 below.

An isolated polynucleotide containing a polymorphic variant nucleotide sequence of the invention may be operably linked to one or more expression regulatory elements in a recombinant expression vector capable of being propagated and expressing the encoded HTR1A protein in a prokaryotic or a eukaryotic host cell. Examples of expression regulatory elements which may be used include, but are not limited to, the lac system, operator and promoter regions of phage lambda, yeast promoters, and promoters derived from vaccinia virus, adenovirus, retroviruses, or SV40. Other regulatory elements include, but are not limited to, appropriate leader sequences, termination codons, polyadenylation signals, and other sequences required for the appropriate transcription and subsequent translation of the nucleic acid sequence in a given host cell. Of course, the correct combinations of expression regulatory elements will depend on the host system used. In addition, it is understood that the expression vector contains any additional elements necessary for its transfer to and subsequent replication in the host cell. Examples of such elements include, but are not limited to, origins of replication and selectable markers. Such expression vectors are commercially available or are readily constructed using methods known to those in the art (e.g., F. Ausubel et al., 1987, in "Current Protocols in Molecular Biology", John Wiley and Sons, New York, New York). Host cells which may be used to express the variant HTR1A sequences of the invention include, but are not limited to, eukaryotic and mammalian cells, such as animal, plant, insect

and yeast cells, and prokaryotic cells, such as E. coli, or algal cells as known in the art. The recombinant expression vector may be introduced into the host cell using any method known to those in the art including, but not limited to, microinjection, electroporation, particle bombardment, transduction, and transfection using DEAE-dextran, lipofection, or calcium phosphate (see e.g., Sambrook et al. (1989) in "Molecular Cloning. A Laboratory Manual", Cold Spring Harbor Press, Plainview, New York). In a preferred aspect, eukaryotic expression vectors that function in eukaryotic cells, and preferably mammalian cells, are used. Non-limiting examples of such vectors include vaccinia virus vectors, adenovirus vectors, herpes virus vectors, and baculovirus transfer vectors. Preferred eukaryotic cell lines include COS cells, CHO cells, HeLa cells, NIH/3T3 cells, and embryonic stem cells (Thomson, J. A. et al., 1998 Science 282:1145-1147). Particularly preferred host cells are mammalian cells.

As will be readily recognized by the skilled artisan, expression of polymorphic variants of the HTR1A gene will produce HTR1A mRNAs varying from each other at any polymorphic site retained in the spliced and processed mRNA molecules. These mRNAs can be used for the preparation of an HTR1A cDNA comprising a nucleotide sequence which is a polymorphic variant of the HTR1A reference coding sequence shown in Figure 4. Thus, the invention also provides HTR1A mRNAs and corresponding cDNAs which comprise a nucleotide sequence that is identical to SEQ ID NO:4 (Fig. 4), or its corresponding RNA sequence, except for having one or both polymorphisms selected from the group consisting of thymine at a position corresponding to nucleotide 423 and guanine at a position corresponding to nucleotide 464, and may also comprise one or more polymorphisms selected from the group consisting of thymine at a position corresponding to nucleotide 47, guanine at a position corresponding to nucleotide 82, adenine at a position corresponding to nucleotide 294, thymine at a position corresponding to nucleotide 552, thymine at a position corresponding to nucleotide 659 and adenine at a position corresponding to nucleotide 818. Fragments of these variant mRNAs and cDNAs are included in the scope of the invention, provided they contain the novel polymorphisms described herein. The invention specifically excludes polynucleotides identical to previously identified and characterized HTR1A cDNAs and fragments thereof. Polynucleotides comprising a variant RNA or DNA sequence may be isolated from a biological sample using well-known molecular biological procedures or may be chemically synthesized.

Genomic and cDNA fragments of the invention comprise at least one novel polymorphic site identified herein and have a length of at least 10 nucleotides and may range up to the full length of the gene. Preferably, a fragment according to the present invention is between 100 and 3000 nucleotides in length, and more preferably between 200 and 2000 nucleotides in length, and most preferably between 500 and 1000 nucleotides in length.

In describing the polymorphic sites identified herein, reference is made to the sense strand of the gene for convenience. However, as recognized by the skilled artisan, nucleic acid molecules containing the HTR1A gene may be complementary double stranded molecules and thus reference to a particular site

on the sense strand refers as well to the corresponding site on the complementary antisense strand. Thus, reference may be made to the same polymorphic site on either strand and an oligonucleotide may be designed to hybridize specifically to either strand at a target region containing the polymorphic site. Thus, the invention also includes single-stranded polynucleotides which are complementary to the sense strand of the HTR1A genomic variants described herein.

Polynucleotides comprising a polymorphic gene variant or fragment may be useful for therapeutic purposes. For example, where a patient could benefit from expression, or increased expression, of a particular HTR1A protein isoform, an expression vector encoding the isoform may be administered to the patient. The patient may be one who lacks the HTR1A isogene encoding that isoform or may already have at least one copy of that isogene.

In other situations, it may be desirable to decrease or block expression of a particular HTR1A isogene. Expression of an HTR1A isogene may be turned off by transforming a targeted organ, tissue or cell population with an expression vector that expresses high levels of untranslatable mRNA for the isogene. Alternatively, oligonucleotides directed against the regulatory regions (e.g., promoter, introns, enhancers, 3′ untranslated region) of the isogene may block transcription. Oligonucleotides targeting the transcription initiation site, e.g., between positions −10 and +10 from the start site are preferred. Similarly, inhibition of transcription can be achieved using oligonucleotides that base-pair with region(s) of the isogene DNA to form triplex DNA (see e.g., Gee et al. in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, N.Y., 1994). Antisense oligonucleotides may also be designed to block translation of HTR1A mRNA transcribed from a particular isogene. It is also contemplated that ribozymes may be designed that can catalyze the specific cleavage of HTR1A mRNA transcribed from a particular isogene.

The oligonucleotides may be delivered to a target cell or tissue by expression from a vector introduced into the cell or tissue *in vivo* or *ex vivo*. Alternatively, the oligonucleotides may be formulated as a pharmaceutical composition for administration to the patient. Oligoribonucleotides and/or oligodeoxynucleotides intended for use as antisense oligonucleotides may be modified to increase stability and half-life. Possible modifications include, but are not limited to phosphorothioate or 2′ O-methyl linkages, and the inclusion of nontraditional bases such as inosine and queosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytosine, guanine, thymine, and uracil which are not as easily recognized by endogenous nucleases.

The invention also provides an isolated polypeptide comprising a polymorphic variant of the reference HTR1A amino acid sequence shown in 5. The location of a variant amino acid in an HTR1A polypeptide or fragment of the invention is identified by aligning its sequence against Fig. 5. An HTR1A protein variant of the invention comprises an amino acid sequence identical to SEQ ID NO:5 except for having glycine at a position corresponding to amino acid position 155, and may also comprise one or more variant amino acids selected from the group consisting of leucine at a position corresponding to

amino acid position 16, valine at a position corresponding to amino acid position 28, leucine at a position corresponding to amino acid position 220 and aspartic acid at a position corresponding to amino acid position 273. The invention specifically excludes amino acid sequences identical to those previously identified for HTR1A, including SEQ ID NO:5, and previously described fragments thereof. HTR1A protein variants included within the invention comprise all amino acid sequences based on SEQ ID NO:5 and having the combination of amino acid variations described in Table 2 below. In preferred embodiments, an HTR1A protein variant of the invention is encoded by an isogene defined by one of the observed haplotypes shown in Table 5.

Table 2. Novel Polymorphic Variants of HTR1A

Polymorphic Variant Number	Amino Acid Position and Identities 16	28	155	220	273
1	P	I	G	R	G
2	P	I	G	R	D
3	P	I	G	L	G
4	P	I	G	L	D
5	P	V	G	R	G
6	P	V	G	R	D
7	P	V	G	L	G
8	P	V	G	L	D
9	L	I	G	R	G
10	L	I	G	R	D
11	L	I	G	L	G
12	L	I	G	L	D
13	L	V	G	R	G
14	L	V	G	R	D
15	L	V	G	L	G
16	L	V	G	L	D

The invention also includes HTR1A peptide variants, which are any fragments of an HTR1A protein variant that contains one or more of the amino acid variations shown in Table 2. An HTR1A peptide variant is at least 6 amino acids in length and is preferably any number between 6 and 30 amino acids long, more preferably between 10 and 25, and most preferably between 15 and 20 amino acids long. Such HTR1A peptide variants may be useful as antigens to generate antibodies specific for one of the above HTR1A isoforms. In addition, the HTR1A peptide variants may be useful in drug screening assays.

An HTR1A variant protein or peptide of the invention may be prepared by chemical synthesis or by expressing one of the variant HTR1A genomic and cDNA sequences as described above. Alternatively, the HTR1A protein variant may be isolated from a biological sample of an individual having an HTR1A isogene which encodes the variant protein. Where the sample contains two different HTR1A isoforms (i.e., the individual has different HTR1A isogenes), a particular HTR1A isoform of the

invention can be isolated by immunoaffinity chromatography using an antibody which specifically binds to that particular HTR1A isoform but does not bind to the other HTR1A isoform.

The expressed or isolated HTR1A protein may be detected by methods known in the art, including Coomassie blue staining, silver staining, and Western blot analysis using antibodies specific for the isoform of the HTR1A protein as discussed further below. HTR1A variant proteins can be purified by standard protein purification procedures known in the art, including differential precipitation, molecular sieve chromatography, ion-exchange chromatography, isoelectric focusing, gel electrophoresis, affinity and immunoaffinity chromatography and the like. (Ausubel et. al., 1987, In Current Protocols in Molecular Biology John Wiley and Sons, New York, New York). In the case of immunoaffinity chromatography, antibodies specific for a particular polymorphic variant may be used.

A polymorphic variant HTR1A gene of the invention may also be fused in frame with a heterologous sequence to encode a chimeric HTR1A protein. The non-HTR1A portion of the chimeric protein may be recognized by a commercially available antibody. In addition, the chimeric protein may also be engineered to contain a cleavage site located between the HTR1A and non-HTR1A portions so that the HTR1A protein may be cleaved and purified away from the non-HTR1A portion.

An additional embodiment of the invention relates to using a novel HTR1A protein isoform in any of a variety of drug screening assays. Such screening assays may be performed to identify agents that bind specifically to all known HTR1A protein isoforms or to only a subset of one or more of these isoforms. The agents may be from chemical compound libraries, peptide libraries and the like. The HTR1A protein or peptide variant may be free in solution or affixed to a solid support. In one embodiment, high throughput screening of compounds for binding to an HTR1A variant may be accomplished using the method described in PCT application WO84/03565, in which large numbers of test compounds are synthesized on a solid substrate, such as plastic pins or some other surface, contacted with the HTR1A protein(s) of interest and then washed. Bound HTR1A protein(s) are then detected using methods well-known in the art.

In another embodiment, a novel HTR1A protein isoform may be used in assays to measure the binding affinities of one or more candidate drugs targeting the HTR1A protein.

In another embodiment, the invention provides antibodies specific for and immunoreactive with one or more of the novel HTR1A variant proteins described herein. The antibodies may be either monoclonal or polyclonal in origin. The HTR1A protein or peptide variant used to generate the antibodies may be from natural or recombinant sources or produced by chemical synthesis using synthesis techniques known in the art. If the HTR1A protein variant is of insufficient size to be antigenic, it may be conjugated, complexed, or otherwise covalently linked to a carrier molecule to enhance the antigenicity of the peptide. Examples of carrier molecules, include, but are not limited to, albumins (e.g., human, bovine, fish, ovine), and keyhole limpet hemocyanin (Basic and Clinical Immunology, 1991, Eds. D.P. Stites, and A.I. Terr, Appleton and Lange, Norwalk Connecticut, San Mateo, California).

In one embodiment, an antibody specifically immunoreactive with one of the novel HTR1A protein isoforms described herein is administered to an individual to neutralize activity of the HTR1A isoform expressed by that individual. The antibody may be formulated as a pharmaceutical composition which includes a pharmaceutically acceptable carrier.

Antibodies specific for and immunoreactive with one of the novel HTR1A protein isoform described herein may be used to immunoprecipitate the HTR1A protein variant from solution as well as react with HTR1A protein isoforms on Western or immunoblots of polyacrylamide gels on membrane supports or substrates. In another preferred embodiment, the antibodies will detect HTR1A protein isoforms in paraffin or frozen tissue sections, or in cells which have been fixed or unfixed and prepared on slides, coverslips, or the like, for use in immunocytochemical, immunohistochemical, and immunofluorescence techniques.

In another embodiment, an antibody specifically immunoreactive with one of the novel HTR1A protein variants described herein is used in immunoassays to detect this variant in biological samples. In this method, an antibody of the present invention is contacted with a biological sample and the formation of a complex between the HTR1A protein variant and the antibody is detected. As described, suitable immunoassays include radioimmunoassay, Western blot assay, immunofluorescent assay, enzyme linked immunoassay (ELISA), chemiluminescent assay, immunohistochemical assay, immunocytochemical assay, and the like (see, e.g., Principles and Practice of Immunoassay, 1991, Eds. Christopher P. Price and David J. Neoman, Stockton Press, New York, New York; Current Protocols in Molecular Biology, 1987, Eds. Ausubel et al., John Wiley and Sons, New York, New York). Standard techniques known in the art for ELISA are described in Methods in Immunodiagnosis, 2nd Ed., Eds. Rose and Bigazzi, John Wiley and Sons, New York 1980; and Campbell et al., 1984, Methods in Immunology, W.A. Benjamin, Inc.). Such assays may be direct, indirect, competitive, or noncompetitive as described in the art (see, e.g., Principles and Practice of Immunoassay, 1991, Eds. Christopher P. Price and David J. Neoman, Stockton Pres, NY, NY; and Oellirich, M., 1984, J. Clin. Chem. Clin. Biochem., 22:895-904). Proteins may be isolated from test specimens and biological samples by conventional methods, as described in Current Protocols in Molecular Biology, supra.

Exemplary antibody molecules for use in the detection and therapy methods of the present invention are intact immunoglobulin molecules, substantially intact immunoglobulin molecules, or those portions of immunoglobulin molecules that contain the antigen binding site. Polyclonal or monoclonal antibodies may be produced by methods conventionally known in the art (e.g., Kohler and Milstein, 1975, Nature, 256:495-497; Campbell Monoclonal Antibody Technology, the Production and Characterization of Rodent and Human Hybridomas, 1985, In: Laboratory Techniques in Biochemistry and Molecular Biology, Eds. Burdon et al., Volume 13, Elsevier Science Publishers, Amsterdam). The antibodies or antigen binding fragments thereof may also be produced by genetic engineering. The

technology for expression of both heavy and light chain genes in E. coli is the subject of PCT patent applications, publication number WO 901443, WO 901443 and WO 9014424 and in Huse et al., 1989, Science, 246:1275-1281. The antibodies may also be humanized (e.g., Queen, C. et al. 1989 Proc. Natl. Acad. Sci. 86;10029).

Effect(s) of the polymorphisms identified herein on expression of HTR1A may be investigated by preparing recombinant cells and/or organisms, preferably recombinant animals, containing a polymorphic variant of the HTR1A gene. As used herein, "expression" includes but is not limited to one or more of the following: transcription of the gene into precursor mRNA; splicing and other processing of the precursor mRNA to produce mature mRNA; mRNA stability; translation of the mature mRNA into HTR1A protein (including codon usage and tRNA availability); and glycosylation and/or other modifications of the translation product, if required for proper expression and function.

To prepare a recombinant cell of the invention, the desired HTR1A isogene may be introduced into the cell in a vector such that the isogene remains extrachromosomal. In such a situation, the gene will be expressed by the cell from the extrachromosomal location. In a preferred embodiment, the HTR1A isogene is introduced into a cell in such a way that it recombines with the endogenous HTR1A gene present in the cell. Such recombination requires the occurrence of a double recombination event, thereby resulting in the desired HTR1A gene polymorphism. Vectors for the introduction of genes both for recombination and for extrachromosomal maintenance are known in the art, and any suitable vector or vector construct may be used in the invention. Methods such as electroporation, particle bombardment, calcium phosphate co-precipitation and viral transduction for introducing DNA into cells are known in the art; therefore, the choice of method may lie with the competence and preference of the skilled practitioner. Examples of cells into which the HTR1A isogene may be introduced include, but are not limited to, continuous culture cells, such as COS, NIH/3T3, and primary or culture cells of the relevant tissue type, i.e., they express the HTR1A isogene. Such recombinant cells can be used to compare the biological activities of the different protein variants.

Recombinant organisms, i.e., transgenic animals, expressing a variant HTR1A gene are prepared using standard procedures known in the art. Preferably, a construct comprising the variant gene is introduced into a nonhuman animal or an ancestor of the animal at an embryonic stage, i.e., the one-cell stage, or generally not later than about the eight-cell stage. Transgenic animals carrying the constructs of the invention can be made by several methods known to those having skill in the art. One method involves transfecting into the embryo a retrovirus constructed to contain one or more insulator elements, a gene or genes of interest, and other components known to those skilled in the art to provide a complete shuttle vector harboring the insulated gene(s) as a transgene, see e.g., U.S. Patent No. 5,610,053. Another method involves directly injecting a transgene into the embryo. A third method involves the use of embryonic stem cells. Examples of animals into which the HTR1A isogenes may be introduced include, but are not limited to, mice, rats, other rodents, and nonhuman primates (see "The Introduction of Foreign

Genes into Mice" and the cited references therein, In: Recombinant DNA, Eds. J.D. Watson, M. Gilman, J. Witkowski, and M. Zoller; W.H. Freeman and Company, New York, pages 254-272). Transgenic animals stably expressing a human HTR1A isogene and producing human HTR1A protein can be used as biological models for studying diseases related to abnormal HTR1A expression and/or activity, and for screening and assaying various candidate drugs, compounds, and treatment regimens to reduce the symptoms or effects of these diseases.

An additional embodiment of the invention relates to pharmaceutical compositions for treating disorders affected by expression or function of a novel HTR1A isogene described herein. The pharmaceutical composition may comprise any of the following active ingredients: a polynucleotide comprising one of these novel HTR1A isogenes; an antisense oligonucleotide directed against one of the novel HTR1A isogenes, a polynucleotide encoding such an antisense oligonucleotide, or another compound which inhibits expression of a novel HTR1A isogene described herein. Preferably, the composition contains the active ingredient in a therapeutically effective amount. By therapeutically effective amount is meant that one or more of the symptoms relating to disorders affected by expression or function of a novel HTR1A isogene is reduced and/or eliminated. The composition also comprises a pharmaceutically acceptable carrier, examples of which include, but are not limited to, saline, buffered saline, dextrose, and water. Those skilled in the art may employ a formulation most suitable for the active ingredient, whether it is a polynucleotide, oligonucleotide, protein, peptide or small molecule antagonist. The pharmaceutical composition may be administered alone or in combination with at least one other agent, such as a stabilizing compound. Administration of the pharmaceutical composition may be by any number of routes including, but not limited to oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, intradermal, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, PA).

For any composition, determination of the therapeutically effective dose of active ingredient and/or the appropriate route of administration is well within the capability of those skilled in the art. For example, the dose can be estimated initially either in cell culture assays or in animal models. The animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans. The exact dosage will be determined by the practitioner, in light of factors relating to the patient requiring treatment, including but not limited to severity of the disease state, general health, age, weight and gender of the patient, diet, time and frequency of administration, other drugs being taken by the patient, and tolerance/response to the treatment.

Information on the identity of genotypes and haplotypes for the HTR1A gene of any particular individual as well as the frequency of such genotypes and haplotypes in any particular population of

individuals is expected to be useful for a variety of basic research and clinical applications. Thus, the invention also provides compositions and methods for detecting the novel HTR1A polymorphisms identified herein.

The compositions comprise at least one HTR1A genotyping oligonucleotide. In one embodiment, an HTR1A genotyping oligonucleotide is a probe or primer capable of hybridizing to a target region that is located close to, or that contains, one of the novel polymorphic sites described herein. As used herein, the term "oligonucleotide" refers to a polynucleotide molecule having less than about 100 nucleotides. A preferred oligonucleotide of the invention is 10 to 35 nucleotides long. More preferably, the oligonucleotide is between 15 and 30, and most preferably, between 20 and 25 nucleotides in length. The oligonucleotide may be comprised of any phosphorylation state of ribonucleotides, deoxyribonucleotides, and acyclic nucleotide derivatives, and other functionally equivalent derivatives. Alternatively, oligonucleotides may have a phosphate-free backbone, which may be comprised of linkages such as carboxymethyl, acetamidate, carbamate, polyamide (peptide nucleic acid (PNA)) and the like (Varma, R. in Molecular Biology and Biotechnology, A Comprehensive Desk Reference, Ed. R. Meyers, VCH Publishers, Inc. (1995), pages 617-620). Oligonucleotides of the invention may be prepared by chemical synthesis using any suitable methodology known in the art, or may be derived from a biological sample, for example, by restriction digestion. The oligonucleotides may be labeled, according to any technique known in the art, including use of radiolabels, fluorescent labels, enzymatic labels, proteins, haptens, antibodies, sequence tags and the like.

Genotyping oligonucleotides of the invention must be capable of specifically hybridizing to a target region of an HTR1A polynucleotide, i.e., an HTR1A isogene. As used herein, specific hybridization means the oligonucleotide forms an anti-parallel double-stranded structure with the target region under certain hybridizing conditions, while failing to form such a structure when incubated with a non-target region or a non-HTR1A polynucleotide under the same hybridizing conditions. Preferably, the oligonucleotide specifically hybridizes to the target region under conventional high stringency conditions. The skilled artisan can readily design and test oligonucleotide probes and primers suitable for detecting polymorphisms in the HTR1A gene using the polymorphism information provided herein in conjunction with the known sequence information for the HTR1A gene and routine techniques.

A nucleic acid molecule such as an oligonucleotide or polynucleotide is said to be a "perfect" or "complete" complement of another nucleic acid molecule if every nucleotide of one of the molecules is complementary to the nucleotide at the corresponding position of the other molecule. A nucleic acid molecule is "substantially complementary" to another molecule if it hybridizes to that molecule with sufficient stability to remain in a duplex form under conventional low-stringency conditions. Conventional hybridization conditions are described, for example, by Sambrook J. et al., in Molecular Cloning, A Laboratory Manual, 2nd Edition, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989) and by Haymes, B.D. et al. in Nucleic Acid Hybridization, A Practical Approach, IRL Press, Washington,

D.C. (1985). While perfectly complementary oligonucleotides are preferred for detecting polymorphisms, departures from complete complementarity are contemplated where such departures do not prevent the molecule from specifically hybridizing to the target region. For example, an oligonucleotide primer may have a non-complementary fragment at its 5′ end, with the remainder of the primer being complementary to the target region. Alternatively, non-complementary nucleotides may be interspersed into the oligonucleotide probe or primer as long as the resulting probe or primer is still capable of specifically hybridizing to the target region.

Preferred genotyping oligonucleotides of the invention are allele-specific oligonucleotides. As used herein, the term allele-specific oligonucleotide (ASO) means an oligonucleotide that is able, under sufficiently stringent conditions, to hybridize specifically to one allele of a gene, or other locus, at a target region containing a polymorphic site while not hybridizing to the corresponding region in another allele(s). As understood by the skilled artisan, allele-specificity will depend upon a variety of readily optimized stringency conditions, including salt and formamide concentrations, as well as temperatures for both the hybridization and washing steps. Examples of hybridization and washing conditions typically used for ASO probes are found in Kogan et al., "Genetic Prediction of Hemophilia A" in PCR Protocols, A Guide to Methods and Applications, Academic Press, 1990 and Ruano et al., 87 Proc. Natl. Acad. Sci. USA 6296-6300, 1990. Typically, an allele-specific oligonucleotide will be perfectly complementary to one allele while containing a single mismatch for another allele.

Allele-specific oligonucleotide probes which usually provide good discrimination between different alleles are those in which a central position of the oligonucleotide probe aligns with the polymorphic site in the target region (e.g., approximately the 7th or 8th position in a 15 mer, the 8th or 9th position in a 16mer, the 10th or 11th position in a 20 mer). A preferred ASO probe for detecting HTR1A gene polymorphisms comprises a nucleotide sequence, listed 5′ to 3′, selected from the group consisting of:

GGAGCGCCTGAAAGC (SEQ ID NO:6) and its complement,
GGAGCGCTTGAAAGC (SEQ ID NO:7) and its complement,
CGGACCCCATCGACT (SEQ ID NO:8) and its complement,
CGGACCCTATCGACT (SEQ ID NO:9) and its complement,
GCCGCTGCGCTCATC (SEQ ID NO:10) and its complement and
GCCGCTGGGCTCATC (SEQ ID NO:11) and its complement.

An allele-specific oligonucleotide primer of the invention has a 3′ terminal nucleotide, or preferably a 3′ penultimate nucleotide, that is complementary to only one nucleotide of a particular SNP, thereby acting as a primer for polymerase-mediated extension only if the allele containing that nucleotide is present. Allele-specific oligonucleotide primers hybridizing to either the coding or noncoding strand are contemplated by the invention. A preferred ASO primer for detecting HTR1A gene polymorphisms comprises a nucleotide sequence, listed 5′ to 3′, selected from the group consisting of:

GCCCAGGGAGCGCCT (SEQ ID NO:12); GGAGCAGCTTTCAGG (SEQ ID NO:13);
GCCCAGGGAGCGCTT (SEQ ID NO:14); GGAGCAGCTTTCAAG (SEQ ID NO:15);
CCATCACGGACCCCA (SEQ ID NO:16); TCACGTAGTCGATGG (SEQ ID NO:17);
CCATCACGGACCCTA (SEQ ID NO:18); TCACGTAGTCGATAGG (SEQ ID NO:19);
CGGCGCGCCGCTGCG (SEQ ID NO:20); GAGCGAGATGAGCGC (SEQ ID NO:21);
CGGCGCGCCGCTGGG (SEQ ID NO:22); and GAGCGAGATGAGCCCA (SEQ ID NO:23).

Other genotyping oligonucleotides of the invention hybridize to a target region located one to several nucleotides downstream of one of the novel polymorphic sites identified herein. Such oligonucleotides are useful in polymerase-mediated primer extension methods for detecting one of the novel polymorphisms described herein and therefore such genotyping oligonucleotides are referred to herein as "primer-extension oligonucleotides". In a preferred embodiment, the 3′-terminus of a primer-extension oligonucleotide is a deoxynucleotide complementary to the nucleotide located immediately adjacent to the polymorphic site. A particularly preferred oligonucleotide primer for detecting HTR1A gene polymorphisms by primer extension terminates in a nucleotide sequence, listed 5′ to 3′, selected from the group consisting of:

CAGGGAGCGC (SEQ ID NO:24); GCAGCTTTCA (SEQ ID NO:25);
TCACGGACCC (SEQ ID NO:26); CGTAGTCGAT (SEQ ID NO:27);
CGCGCCGCTG (SEQ ID NO:28); and CGAGATGAGC (SEQ ID NO:29).

In some embodiments, a composition contains two or more differently labeled genotyping oligonucleotides for simultaneously probing the identity of nucleotides at two or more polymorphic sites. It is also contemplated that primer compositions may contain two or more sets of allele-specific primer pairs to allow simultaneous targeting and amplification of two or more regions containing a polymorphic site.

HTR1A genotyping oligonucleotides of the invention may also be immobilized on or synthesized on a solid surface such as a microchip, bead, or glass slide (see, e.g., WO 98/20020 and WO 98/20019). Such immobilized genotyping oligonucleotides may be used in a variety of polymorphism detection assays, including but not limited to probe hybridization and polymerase extension assays. Immobilized HTR1A genotyping oligonucleotides of the invention may comprise an ordered array of oligonucleotides designed to rapidly screen a DNA sample for polymorphisms in multiple genes at the same time.

In another embodiment, the invention provides a kit comprising at least two genotyping oligonucleotides packaged in separate containers. The kit may also contain other components such as

hybridization buffer (where the oligonucleotides are to be used as a probe) packaged in a separate container. Alternatively, where the oligonucleotides are to be used to amplify a target region, the kit may contain, packaged in separate containers, a polymerase and a reaction buffer optimized for primer extension mediated by the polymerase, such as PCR.

The above described oligonucleotide compositions and kits are useful in methods for genotyping and/or haplotyping the HTR1A gene in an individual. As used herein, the terms "HTR1A genotype" and "HTR1A haplotype" mean the genotype or haplotype contains the nucleotide pair or nucleotide, respectively, that is present at one or more of the novel polymorphic sites described herein and may optionally also include the nucleotide pair or nucleotide present at one or more additional polymorphic sites in the HTR1A gene. The additional polymorphic sites may be currently known polymorphic sites or sites that are subsequently discovered.

One embodiment of the genotyping method involves isolating from the individual a nucleic acid mixture comprising the two copies of the HTR1A gene, or a fragment thereof, that are present in the individual, and determining the identity of the nucleotide pair at one or more of the polymorphic sites selected from PS1, PS5 and PS6 in the two copies to assign an HTR1A genotype to the individual. As will be readily understood by the skilled artisan, the two "copies" of a gene in an individual may be the same allele or may be different alleles. In a preferred embodiment of the genotyping method, the identity of the nucleotide pair atone or more of the polymorphic sites selected from the group consisting of PS3, PS4 and PS8 is also determined. In a particularly preferred embodiment, the genotyping method comprises determining the identity of the nucleotide pair at each of PS1-9.

Typically, the nucleic acid mixture is isolated from a biological sample taken from the individual, such as a blood sample or tissue sample. Suitable tissue samples include whole blood, semen saliva, tears, urine, fecal material, sweat, buccal, skin and hair. The nucleic acid mixture may be comprised of genomic DNA, mRNA, or cDNA and, in the latter two cases, the biological sample must be obtained from an organ in which the HTR1A gene is expressed. Furthermore it will be understood by the skilled artisan that mRNA or cDNA preparations would not be used to detect polymorphisms located in introns or in 5′ and 3′ nontranscribed regions. If an HTR1A gene fragment is isolated, it must contain the polymorphic site(s) to be genotyped.

One embodiment of the haplotyping method comprises isolating from the individual a nucleic acid molecule containing only one of the two copies of the HTR1A gene, or a fragment thereof, that is present in the individual and determining in that copy the identity of the nucleotide at one or more of the polymorphic sites PS1, PS5 and PS6 in that copy to assign an HTR1A haplotype to the individual. The nucleic acid may be isolated using any method capable of separating the two copies of the HTR1A gene or fragment such as one of the methods described above for preparing HTR1A isogenes, with targeted *in vivo* cloning being the preferred approach. As will be readily appreciated by those skilled in the art, any individual clone will only provide haplotype information on one of the two HTR1A gene copies present

in an individual. If haplotype information is desired for the individual's other copy, additional HTR1A clones will need to be examined. Typically, at least five clones should be examined to have more than a 90% probability of haplotyping both copies of the HTR1A gene in an individual. In some embodiments, the haplotyping method also comprises identifying the nucleotide at one or more of the polymorphic sites PS2, PS3, PS4 , PS7, PS8 andn PS9. In a particularly preferred embodiment, the nucleotide at each of PS1-9 is identified.

In a preferred embodiment, an HTR1A haplotype pair is determined for an individual by identifying the phased sequence of nucleotides at one or more of the polymorphic sites selected from PS1, PS5 and PS6 in each copy of the HTR1A gene that is present in the individual. In a particularly preferred embodiment, the haplotyping method comprises identifying the phased sequence of nucleotides at each of PS1-9 in each copy of the HTR1A gene. When haplotyping both copies of the gene, the identifying step is preferably performed with each copy of the gene being placed in separate containers. However, it is also envisioned that if the two copies are labeled with different tags, or are otherwise separately distinguishable or identifiable, it could be possible in some cases to perform the method in the same container. For example, if first and second copies of the gene are labeled with different first and second fluorescent dyes, respectively, and an allele-specific oligonucleotide labeled with yet a third different fluorescent dye is used to assay the polymorphic site(s), then detecting a combination of the first and third dyes would identify the polymorphism in the first gene copy while detecting a combination of the second and third dyes would identify the polymorphism in the second gene copy.

In both the genotyping and haplotyping methods, the identity of a nucleotide (or nucleotide pair) at a polymorphic site(s) may be determined by amplifying a target region(s) containing the polymorphic site(s) directly from one or both copies of the HTR1A gene, or fragment thereof, and the sequence of the amplified region(s) determined by conventional methods. It will be readily appreciated by the skilled artisan that only one nucleotide will be detected at a polymorphic site in individuals who are homozygous at that site, while two different nucleotides will be detected if the individual is heterozygous for that site. The polymorphism may be identified directly, known as positive-type identification, or by inference, referred to as negative-type identification. For example, where a SNP is known to be guanine and cytosine in a reference population, a site may be positively determined to be either guanine or cytosine for an individual homozygous at that site, or both guanine and cytosine, if the individual is heterozygous at that site. Alternatively, the site may be negatively determined to be not guanine (and thus cytosine/cytosine) or not cytosine (and thus guanine/guanine).

In addition, the identity of the allele(s) present at any of the novel polymorphic sites described herein may be indirectly determined by genotyping a polymorphic site not disclosed herein that is in linkage disequilibrium with the polymorphic site that is of interest. Two sites are said to be in linkage disequilibrium if the presence of a particular variant at one site enhances the predictability of another variant at the second site (Stevens, JC 1999, *Mol. Diag.* 4: 309-17). Polymorphic sites in linkage

disequilibrium with the presently disclosed polymorphic sites may be located in regions of the gene or in other genomic regions not examined herein. Genotyping of a polymorphic site in linkage disequilibrium with the novel polymorphic sites described herein may be performed by, but is not limited to, any of the above-mentioned methods for detecting the identity of the allele at a polymorphic site.

The target region(s) may be amplified using any oligonucleotide-directed amplification method, including but not limited to polymerase chain reaction (PCR) (U.S. Patent No. 4,965,188), ligase chain reaction (LCR) (Barany et al., Proc. Natl. Acad. Sci. USA 88:189-193, 1991; WO90/01069), and oligonucleotide ligation assay (OLA) (Landegren et al., Science 241:1077-1080, 1988). Oligonucleotides useful as primers or probes in such methods should specifically hybridize to a region of the nucleic acid that contains or is adjacent to the polymorphic site. Typically, the oligonucleotides are between 10 and 35 nucleotides in length and preferably, between 15 and 30 nucleotides in length. Most preferably, the oligonucleotides are 20 to 25 nucleotides long. The exact length of the oligonucleotide will depend on many factors that are routinely considered and practiced by the skilled artisan.

Other known nucleic acid amplification procedures may be used to amplify the target region including transcription-based amplification systems (U.S. Patent No. 5,130,238; EP 329,822; U.S. Patent No. 5,169,766, WO89/06700) and isothermal methods (Walker et al., *Proc. Natl. Acad. Sci. USA* **89**:392-396, 1992).

A polymorphism in the target region may also be assayed before or after amplification using one of several hybridization-based methods known in the art. Typically, allele-specific oligonucleotides are utilized in performing such methods. The allele-specific oligonucleotides may be used as differently labeled probe pairs, with one member of the pair showing a perfect match to one variant of a target sequence and the other member showing a perfect match to a different variant. In some embodiments, more than one polymorphic site may be detected at once using a set of allele-specific oligonucleotides or oligonucleotide pairs. Preferably, the members of the set have melting temperatures within 5°C, and more preferably within 2°C, of each other when hybridizing to each of the polymorphic sites being detected.

Hybridization of an allele-specific oligonucleotide to a target polynucleotide may be performed with both entities in solution, or such hybridization may be performed when either the oligonucleotide or the target polynucleotide is covalently or noncovalently affixed to a solid support. Attachment may be mediated, for example, by antibody-antigen interactions, poly-L-Lys, streptavidin or avidin-biotin, salt bridges, hydrophobic interactions, chemical linkages, UV cross-linking baking, etc. Allele-specific oligonucleotides may be synthesized directly on the solid support or attached to the solid support subsequent to synthesis. Solid-supports suitable for use in detection methods of the invention include substrates made of silicon, glass, plastic, paper and the like, which may be formed, for example, into wells (as in 96-well plates), slides, sheets, membranes, fibers, chips, dishes, and beads. The solid support may be treated, coated or derivatized to facilitate the immobilization of the allele-specific oligonucleotide

or target nucleic acid.

The genotype or haplotype for the HTR1A gene of an individual may also be determined by hybridization of a nucleic sample containing one or both copies of the gene to nucleic acid arrays and subarrays such as described in WO 95/11995. The arrays would contain a battery of allele-specific oligonucleotides representing each of the polymorphic sites to be included in the genotype or haplotype.

The identity of polymorphisms may also be determined using a mismatch detection technique, including but not limited to the RNase protection method using riboprobes (Winter et al., Proc. Natl. Acad. Sci. USA 82:7575, 1985; Meyers et al., Science 230:1242, 1985) and proteins which recognize nucleotide mismatches, such as the E. coli mutS protein (Modrich, P. Ann. Rev. Genet. 25:229-253, 1991). Alternatively, variant alleles can be identified by single strand conformation polymorphism (SSCP) analysis (Orita et al., Genomics 5:874-879, 1989; Humphries et al., in Molecular Diagnosis of Genetic Diseases, R. Elles, ed., pp. 321-340, 1996) or denaturing gradient gel electrophoresis (DGGE) (Wartell et al., Nucl. Acids Res. 18:2699-2706, 1990; Sheffield et al., Proc. Natl. Acad. Sci. USA 86:232-236, 1989).

A polymerase-mediated primer extension method may also be used to identify the polymorphism(s). Several such methods have been described in the patent and scientific literature and include the "Genetic Bit Analysis" method (WO92/15712) and the ligase/polymerase mediated genetic bit analysis (U.S. Patent 5,679,524. Related methods are disclosed in WO91/02087, WO90/09455, WO95/17676, U.S. Patent Nos. 5,302,509, and 5,945,283. Extended primers containing a polymorphism may be detected by mass spectrometry as described in U.S. Patent No. 5,605,798. Another primer extension method is allele-specific PCR (Ruaño et al., Nucl. Acids Res. 17:8392, 1989; Ruaño et al., Nucl. Acids Res. 19, 6877-6882, 1991; WO 93/22456; Turki et al., J. Clin. Invest. 95:1635-1641, 1995). In addition, multiple polymorphic sites may be investigated by simultaneously amplifying multiple regions of the nucleic acid using sets of allele-specific primers as described in Wallace et al. (WO89/10414).

In another aspect of the invention, an individual's HTR1A haplotype pair is predicted from its HTR1A genotype using information on haplotype pairs known to exist in a reference population. In its broadest embodiment, the haplotyping prediction method comprises identifying an HTR1A genotype for the individual at two or more polymorphic sites selected from PS1, PS5 and PS6, enumerating all possible haplotype pairs which are consistent with the genotype, accessing data containing HTR1A haplotype pairs identified in a reference population, and assigning a haplotype pair to the individual that is consistent with the data. In one embodiment, the reference haplotype pairs include the HTR1A haplotype pairs shown in Table 4.

Generally, the reference population should be composed of randomly-selected individuals representing the major ethnogeographic groups of the world. A preferred reference population for use in the methods of the present invention comprises an approximately equal number of individuals from

Caucasian, African American, Asian and Hispanic-Latino population groups with the minimum number of each group being chosen based on how rare a haplotype one wants to be guaranteed to see. For example, if one wants to have a q% chance of not missing a haplotype that exists in the population at a p% frequency of occurring in the reference population, the number of individuals (n) who must be sampled is given by $2n=\log(1-q)/\log(1-p)$ where p and q are expressed as fractions. A preferred reference population allows the detection of any haplotype whose frequency is at least 10% with about 99% certainty and comprises about 20 unrelated individuals from each of the four population groups named above. A particularly preferred reference population includes a 3-generation family representing one or more of the four population groups to serve as controls for checking quality of haplotyping procedures.

In a preferred embodiment, the haplotype frequency data for each ethnogeographic group is examined to determine whether it is consistent with Hardy-Weinberg equilibrium. Hardy-Weinberg equilibrium (D.L. Hartl et al., Principles of Population Genomics, Sinauer Associates (Sunderland, MA), 3rd Ed., 1997) postulates that the frequency of finding the haplotype pair H_1 / H_2 is equal to $p_{H-W}(H_1 / H_2) = 2p(H_1)p(H_2)$ if $H_1 \neq H_2$ and $p_{H-W}(H_1 / H_2) = p(H_1)p(H_2)$ if $H_1 = H_2$. A statistically significant difference between the observed and expected haplotype frequencies could be due to one or more factors including significant inbreeding in the population group, strong selective pressure on the gene, sampling bias, and/or errors in the genotyping process. If large deviations from Hardy-Weinberg equilibrium are observed in an ethnogeographic group, the number of individuals in that group can be increased to see if the deviation is due to a sampling bias. If a larger sample size does not reduce the difference between observed and expected haplotype pair frequencies, then one may wish to consider haplotyping the individual using a direct haplotyping method such as, for example, CLASPER System™ technology (U.S. Patent No. 5,866,404), SMD, or allele-specific long-range PCR (Michalotos-Beloin et al., Nucleic Acids Res. 24:4841-4843, 1996).

In one embodiment of this method for predicting an HTR1A haplotype pair, the assigning step involves performing the following analysis. First, each of the possible haplotype pairs is compared to the haplotype pairs in the reference population. Generally, only one of the haplotype pairs in the reference population matches a possible haplotype pair and that pair is assigned to the individual. Occasionally, only one haplotype represented in the reference haplotype pairs is consistent with a possible haplotype pair for an individual, and in such cases the individual is assigned a haplotype pair containing this known haplotype and a new haplotype derived by subtracting the known haplotype from the possible haplotype pair. In rare cases, either no haplotypes in the reference population are consistent with the possible haplotype pairs, or alternatively, multiple reference haplotype pairs are consistent with the possible haplotype pairs. In such cases, the individual is preferably haplotyped using a direct molecular haplotyping method such as, for example, CLASPER System™ technology (U.S. Patent No. 5,866,404), SMD, or allele-specific long-range PCR (Michalotos-Beloin et al., Nucleic Acids Res. 24:4841-4843,

1996).

The invention also provides a method for determining the frequency of an HTR1A genotype or HTR1A haplotype in a population. The method comprises determining the genotype or the haplotype pair for the HTR1A gene that is present in each member of the population, wherein the genotype or haplotype comprises the nucleotide pair or nucleotide detected at one or more of the polymorphic sites PS1, PS5 and PS6 in the HTR1A gene; and calculating the frequency any particular genotype or haplotype is found in the population. The population may be a reference population, a family population, a same sex population, a population group, a trait population (e.g., a group of individuals exhibiting a trait of interest such as a medical condition or response to a therapeutic treatment).

In another aspect of the invention, frequency data for HTR1A genotypes and/or haplotypes found in a reference population are used in a method for identifying an association between a trait and an HTR1A genotype or an HTR1A haplotype. The trait may be any detectable phenotype, including but not limited to susceptibility to a disease or response to a treatment. The method involves obtaining data on the frequency of the genotype(s) or haplotype(s) of interest in a reference population as well as in a population exhibiting the trait. Frequency data for one or both of the reference and trait populations may be obtained by genotyping or haplotyping each individual in the populations using one of the methods described above. The haplotypes for the trait population may be determined directly or, alternatively, by the predictive genotype to haplotype approach described above. In another embodiment, the frequency data for the reference and/or trait populations is obtained by accessing previously determined frequency data, which may be in written or electronic form. For example, the frequency data may be present in a database that is accessible by a computer. Once the frequency data is obtained, the frequencies of the genotype(s) or haplotype(s) of interest in the reference and trait populations are compared. In a preferred embodiment, the frequencies of all genotypes and/or haplotypes observed in the populations are compared. If a particular genotype or haplotype for the HTR1A gene is more frequent in the trait population than in the reference population at a statistically significant amount, then the trait is predicted to be associated with that HTR1A genotype or haplotype. Preferably, the HTR1A genotype or haplotype being compared in the trait and reference populations is selected from the full-genotypes and full-haplotypes shown in Tables 4 and 5, respectively, or from sub-genotypes and sub-haplotypes derived from these genotypes and haplotypes.

In a preferred embodiment of the method, the trait of interest is a clinical response exhibited by a patient to some therapeutic treatment, for example, response to a drug targeting HTR1A or response to a therapeutic treatment for a medical condition. As used herein, "medical condition" includes but is not limited to any condition or disease manifested as one or more physical and/or psychological symptoms for which treatment is desirable, and includes previously and newly identified diseases and other disorders. As used herein the term "clinical response" means any or all of the following: a quantitative measure of the response, no response, and adverse response (i.e., side effects).

In order to deduce a correlation between clinical response to a treatment and an HTR1A genotype or haplotype, it is necessary to obtain data on the clinical responses exhibited by a population of individuals who received the treatment, hereinafter the "clinical population". This clinical data may be obtained by analyzing the results of a clinical trial that has already been run and/or the clinical data may be obtained by designing and carrying out one or more new clinical trials. As used herein, the term "clinical trial" means any research study designed to collect clinical data on responses to a particular treatment, and includes but is not limited to phase I, phase II and phase III clinical trials. Standard methods are used to define the patient population and to enroll subjects.

It is preferred that the individuals included in the clinical population have been graded for the existence of the medical condition of interest. This is important in cases where the symptom(s) being presented by the patients can be caused by more than one underlying condition, and where treatment of the underlying conditions are not the same. An example of this would be where patients experience breathing difficulties that are due to either asthma or respiratory infections. If both sets were treated with an asthma medication, there would be a spurious group of apparent non-responders that did not actually have asthma. These people would affect the ability to detect any correlation between haplotype and treatment outcome. This grading of potential patients could employ a standard physical exam or one or more lab tests. Alternatively, grading of patients could use haplotyping for situations where there is a strong correlation between haplotype pair and disease susceptibility or severity.

The therapeutic treatment of interest is administered to each individual in the trial population and each individual's response to the treatment is measured using one or more predetermined criteria. It is contemplated that in many cases, the trial population will exhibit a range of responses and that the investigator will choose the number of responder groups (e.g., low, medium, high) made up by the various responses. In addition, the HTR1A gene for each individual in the trial population is genotyped and/or haplotyped, which may be done before or after administering the treatment.

After both the clinical and polymorphism data have been obtained, correlations between individual response and HTR1A genotype or haplotype content are created. Correlations may be produced in several ways. In one method, individuals are grouped by their HTR1A genotype or haplotype (or haplotype pair) (also referred to as a polymorphism group), and then the averages and standard deviations of clinical responses exhibited by the members of each polymorphism group are calculated.

These results are then analyzed to determine if any observed variation in clinical response between polymorphism groups is statistically significant. Statistical analysis methods which may be used are described in L.D. Fisher and G. vanBelle, "Biostatistics: A Methodology for the Health Sciences", Wiley-Interscience (New York) 1993. This analysis may also include a regression calculation of which polymorphic sites in the PTGS2 gene give the most significant contribution to the differences in phenotype. One regression model useful in the invention is described in the PCT Application entitled

"Methods for Obtaining and Using Haplotype Data", filed June 26, 2000.

A second method for finding correlations between HTR1A haplotype content and clinical responses uses predictive models based on error-minimizing optimization algorithms. One of many possible optimization algorithms is a genetic algorithm (R. Judson, "Genetic Algorithms and Their Uses in Chemistry" in Reviews in Computational Chemistry, Vol. 10, pp. 1-73, K. B. Lipkowitz and D. B. Boyd, eds. (VCH Publishers, New York, 1997). Simulated annealing (Press et al., "Numerical Recipes in C: The Art of Scientific Computing", Cambridge University Press (Cambridge) 1992, Ch. 10), neural networks (E. Rich and K. Knight, "Artificial Intelligence", 2nd Edition (McGraw-Hill, New York, 1991, Ch. 18), standard gradient descent methods (Press et al., supra Ch. 10), or other global or local optimization approaches (see discussion in Judson, supra) could also be used. Preferably, the correlation is found using a genetic algorithm approach as described in PCT Application entitled "Methods for Obtaining and Using Haplotype Data", filed June 26, 2000.

Correlations may also be analyzed using analysis of variation (ANOVA) techniques to determine how much of the variation in the clinical data is explained by different subsets of the polymorphic sites in the PTGS2 gene. As described in PCT Application entitled "Methods for Obtaining and Using Haplotype Data", filed June 26, 2000, ANOVA is used to test hypotheses about whether a response variable is caused by or correlated with one or more traits or variables that can be measured (Fisher and vanBelle, supra, Ch. 10).

From the analyses described above, a mathematical model may be readily constructed by the skilled artisan that predicts clinical response as a function of HTR1A genotype or haplotype content. Preferably, the model is validated in one or more follow-up clinical trials designed to test the model.

The identification of an association between a clinical response and a genotype or haplotype (or haplotype pair) for the HTR1A gene may be the basis for designing a diagnostic method to determine those individuals who will or will not respond to the treatment, or alternatively, will respond at a lower level and thus may require more treatment, i.e., a greater dose of a drug. The diagnostic method may take one of several forms: for example, a direct DNA test (i.e., genotyping or haplotyping one or more of the polymorphic sites in the HTR1A gene), a serological test, or a physical exam measurement. The only requirement is that there be a good correlation between the diagnostic test results and the underlying HTR1A genotype or haplotype that is in turn correlated with the clinical response. In a preferred embodiment, this diagnostic method uses the predictive haplotyping method described above.

Any or all analytical and mathematical operations involved in practicing the methods of the present invention may be implemented by a computer. In addition, the computer may execute a program that generates views (or screens) displayed on a display device and with which the user can interact to view and analyze large amounts of information relating to the HTR1A gene and its genomic variation, including chromosome location, gene structure, and gene family, gene expression data, polymorphism data, genetic sequence data, and clinical data population data (e.g., data on ethnogeographic origin,

clinical responses, genotypes, and haplotypes for one or more populations). The HTR1A polymorphism data described herein may be stored as part of a relational database (e.g., an instance of an Oracle database or a set of ASCII flat files). These polymorphism data may be stored on the computer's hard drive or may, for example, be stored on a CD ROM or on one or more other storage devices accessible by the computer. For example, the data may be stored on one or more databases in communication with the computer via a network.

Preferred embodiments of the invention are described in the following examples. Other embodiments within the scope of the claims herein will be apparent to one skilled in the art from consideration of the specification or practice of the invention as disclosed herein. It is intended that the specification, together with the examples, be considered exemplary only, with the scope and spirit of the invention being indicated by the claims which follow the examples.

EXAMPLES

The Examples herein are meant to exemplify the various aspects of carrying out the invention and are not intended to limit the scope of the invention in any way. The Examples do not include detailed descriptions for conventional methods employed, such as in the performance of genomic DNA isolation, PCR and sequencing procedures. Such methods are well-known to those skilled in the art and are described in numerous publications, for example, Sambrook, Fritsch, and Maniatis, "Molecular Cloning: A Laboratory Manual", 2nd Edition, Cold Spring Harbor Laboratory Press, USA, (1989).

Example 1

This example illustrates examination of various regions of the HTR1A gene for polymorphic sites.

Amplification of Target Regions

The following target regions of the HTR1A gene were amplified using the PCR primer pairs listed below, with the sequences presented in the 5′ to 3′ direction and nucleotide positions shown for each region corresponding to the indicated GenBank Accession No.

```
Accession Number: Z11168
Fragment 1
Forward Primer
678-703                    AAGAGGCAGAAGAGAGAGAAGAGAGG (SEQ ID NO:30)
Reverse Primer
Complement of 1081-1060    AAGTTTCGGAGGAAGGGAATGC (SEQ ID NO:31)
PCR product  404 nt

Accession Number: M83181
Fragment 2
Forward Primer
215-236          AAAGCTGCTCCTCGGAGATACC (SEQ ID NO:32)
```

Reverse Primer
Complement of 774-753 CACAGGTGCAAGATGGATGAGG (SEQ ID NO:33)
PCR product 560 nt

Fragment 3
Forward Primer
463-487 CACTACTGGTATCTCCGACGTGACC (SEQ ID NO:34)
Reverse Primer
Complement of 1049-1027 GGAATATGCGCCCATAGAGAACC (SEQ ID NO:35)
PCR product 587 nt

Fragment 4
Forward Primer
747-767 GCTGCACCTCATCCATCTTGC (SEQ ID NO:36)
Reverse Primer
Complement of 1309-1287 AGGCAAGTGCTCTTTGGAGTTGC (SEQ ID NO:37)
PCR product 563 nt

Fragment 5
Forward Primer
1016-1038 CTGCTCATGCTGGTTCTCTATGG (SEQ ID NO:38)
Reverse Primer
Complement of 1584-1563 ACGGGGTTAAGCAGAGAGTTGG (SEQ ID NO:39)
PCR product 569 nt

Fragment 6
Forward Primer
1280-1301 CGAGTGGGCAACTCCAAAGAGC (SEQ ID NO:40)
Reverse Primer
Complement of 1807-1788 AGGGGACCAGGTCTGCAAGC (SEQ ID NO:41)
PCR product 528 nt

These primer pairs were used in PCR reactions containing genomic DNA isolated from immortalized cell lines for each member of the Index Repository. The PCR reactions were carried out under the following conditions:

Reaction volume	= 20 µl
10 x Advantage 2 Polymerase reaction buffer (Clontech)	= 2 µl
100 ng of human genomic DNA	= 1 µl
10 mM dNTP	= 0.4 µl
Advantage 2 Polymerase enzyme mix (Clontech)	= 0.2 µl
Forward Primer (10 µM)	= 0.4 µl
Reverse Primer (10 µM)	= 0.4 µl
Water	=15.6µl

Amplification profile:
94°C - 2 min. 1 cycle

94°C - 30 sec.
70°C - 45 sec.
72°C - 1 min. } 10 cycles

94°C - 30 sec.
64°C - 45 sec.
72°C - 1 min. } 35 cycles

Sequencing of PCR Products

The PCR products were purified by Solid Phase Reversible Immobilization using the protocol developed by the Whitehead Genome Center. A detailed protocol can be found at http://www.genome.wi.mit.edu/sequencing/protocols/pure/SPRI_pcr.html.

Briefly, five µl of carboxyl coated magnetic beads (10 mg/ml) and 60 µl of HYB BUFFER (2.5M NaCl/20% PEG 8000) were added to each PCR reaction mixture (20 µl). The reaction mixture was mixed well and incubated at room temperature (RT) for 10 min. The microtitre plate was placed on a magnet for 2 min and the beads washed twice with 150 µl of 70% EtOH. The beads were air dried for 2 min and the DNA was eluted in 25 µl of distilled water and incubated at RT for 5 min. The beads were magnetically separated and the supernatant removed for testing and sequencing.

The purified PCR products were sequenced in both directions using the primer sets described previously or those listed, in the 5′ to 3′ direction, below.

Accession Number: Z11168
Fragment 1
Forward Primer
710-729 GAGGGGGAGAGAGGGAAGGA (SEQ ID NO:42)
Reverse Primer
Complement of 1062-1043 TGCAGAGACCCAAGCAGGAA (SEQ ID NO:43)

Accession Number: M83181
Fragment 2
Forward Primer
267-286 GGGTCTCTGCATTCCCTTCC (SEQ ID NO:44)
Reverse Primer
Complement of 731-712 CGATGAACAGGTCGCAGGTT (SEQ ID NO:45)

Fragment 3
Forward Primer
518-537 CTGGGCACGCTCATCTTCTG (SEQ ID NO:46)
Reverse Primer
Complement of 1023-1004 ATGAGCAGCAGCGGGATGTA (SEQ ID NO:47)

Fragment 4
Forward Primer
789-808 ACAGGTACTGGGCCATCACG (SEQ ID NO:48)
Reverse Primer
Complement of 1279-1260 GTGCACCTCGATCACCTCCA (SEQ ID NO:49)

Fragment 5
Forward Primer
1065-1082 GCATCCGCAAGACGGTCA (SEQ ID NO:50)
Reverse Primer

Complement of 1548-1529 TTGATTATGGCGCCCAACAG (SEQ ID NO:51)

Fragment 6
Forward Primer
1322-1341 GCTGGTCCTACCCCTTGTGC (SEQ ID NO:52)
Reverse Primer
Complement of 1787-1768 CGTGAGCGGAGCAGAGAGAA (SEQ ID NO:53)

Analysis of Sequences for Polymorphic Sites

Sequences were analyzed for the presence of polymorphisms using the Polyphred program (Nickerson et al., Nucleic Acids Res. 14:2745-2751, 1997). The presence of a polymorphism was confirmed on both strands. The polymorphisms and their locations in the HTR1A gene are listed in Table 3 below.

Table 3. Polymorphic Sites Identified in the HTR1A Gene

Polymorphic Site Number	Nucleotide Position in Indicated Acc. #	Nucleotide Position in Figure 3	Reference Allele	Variant Allele
PS1	997(Acc#Z11168)	996	C	T
PS2 [R]	438(Acc#M83181)	1222	C	T
PS3 [R]	473(Acc#M83181)	1257	A	G
PS4 [R]	685(Acc#M83181)	1469	G	A
PS5	814(Acc#M83181)	1598	C	T
PS6	855(Acc#M83181)	1639	C	G
PS7 [R]	943(Acc#M83181)	1727	C	T
PS8 [R]	1050(Acc#M83181)	1834	G	T
PS9 [R]	1209(Acc#M83181)	1993	G	A

[R] Reported in previous literature

Example 2

This example illustrates analysis of the HTR1A polymorphisms identified in the Index Repository for human genotypes and haplotypes.

The different genotypes containing these polymorphisms that were observed in the reference population are shown in Table 4 below, with the haplotype pair indicating the combination of haplotypes determined for the individual using the haplotype derivation protocol described below. In Table 4, homozygous positions are indicated by one nucleotide and heterozygous positions are indicated by two nucleotides Missing nucleotides in any given genotype in Table 4 can typically be inferred based on linkage disequilibrium and/or Mendelian inheritance.

Table 4. Genotypes and Haplotype Pairs Observed for HTR1A Gene

Genotype Number	Polymorphic Sites*										
	PS1	PS2	PS3	PS4	PS5	PS6	PS7	PS8	PS9	HAP	Pair
1	C	C	A	G	C	C	C	G	G	3	3
2	C	C	A	A	C	C	C	G	G	1	1
3	C	C	A	G	C	C	C	G/T	G	3	4
4	C	C	A/G	G	C	C	C	G	G	3	8
5	C/T	C	A	G/A	C	C	C	G	G	3	2
6	C	C	A	G	C/T	C	C	G	G	3	7
7	C	C	A	G/A	C	C	C	G	G	3	1
8	C	C/T	A	G	C	C	C	G	G/A	3	9
9	C	C	A	G	C	C/G	C	G	G	3	6
10	C	C	A	G	C	C	C/T	G	G	3	5
11	C	C/T	A	G	C	C	C	G	G	3	10

* homozygous positions indicated by one nucleotide; heterozugous positions indicated by two nucleotides

The haplotype pairs shown in Table 4 were estimated from the unphased genotypes using an extension of Clark's algorithm (Clark, A.G. (1990) *Mol Bio Evol* 7, 111-122), as described in U.S. Provisional Patent Application filed April 19, 2000 and entitled "A Method and System for Determining Haplotypes from a Collection of Polymorphisms". In this method, haplotypes are assigned directly from individuals who are homozygous at all sites or heterozygous at no more than one of the variable sites. This list of haplotypes is then used to deconvolute the unphased genotypes in the remaining (multiply heterozygous) individuals.

By following this protocol, it was determined that the Index Repository examined herein and, by extension, the general population contains the 10 human HTR1A haplotypes shown in Table 5 below.

Table 5. Haplotypes Identified in the HTR1A Gene

Haplotype Number	Polymorphic Sites								
	PS 1	PS 2	PS 3	PS 4	PS 5	PS 6	PS 7	PS 8	PS 9
1	C	C	A	A	C	C	C	G	G
2	T	C	A	A	C	C	C	G	G
3	C	C	A	G	C	C	C	G	G
4	C	C	A	G	C	C	C	T	G
5	C	C	A	G	C	C	T	G	G
6	C	C	A	G	C	G	C	G	G
7	C	C	A	G	T	C	C	G	G
8	C	C	G	G	C	C	C	G	G
9	C	T	A	G	C	C	C	G	A
10	C	T	A	G	C	C	C	G	G

In view of the above, it will be seen that the several advantages of the invention are achieved

and other advantageous results attained.

As various changes could be made in the above methods and compositions without departing from the scope of the invention, it is intended that all matter contained in the above description and shown in the accompanying drawings shall be interpreted as illustrative and not in a limiting sense.

All references cited in this specification, including patents and patent applications, are hereby incorporated in their entirety by reference. The discussion of references herein is intended merely to summarize the assertions made by their authors and no admission is made that any reference constitutes prior art. Applicants reserve the right to challenge the accuracy and pertinency of the cited references.

What is Claimed is:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of:
   (a) a first nucleotide sequence which is a polymorphic variant of a reference sequence for 5-hydroxytryptamine receptor 1A gene (HTR1A) gene or a fragment thereof, wherein the reference sequence comprises SEQ ID NO:3, and the polymorphic variant comprises at least one polymorphism selected from the group consisting of thymine at PS1, thymine at PS5 and guanine at PS6; and
   (b) a second nucleotide sequence which is complementary to the first nucleotide sequence.
2. The isolated polynucleotide of claim 1 which comprises an HTR1A isogene.
3. The isolated polynucleotide of claim 1 which is a DNA molecule and comprises both the first and second nucleotide sequences and further comprises expression regulatory elements operably linked to the first nucleotide sequence.
4. A recombinant organism transformed or transfected with the isolated polynucleotide of claim 1, wherein the organism expresses an HTR1A protein encoded by the first nucleotide sequence.
5. The recombinant organism of claim 4 which is a nonhuman transgenic animal.
6. The isolated polynucleotide of claim 1, wherein the first nucleotide sequence is a polymorphic variant of a fragment of the HTR1A gene, the fragment comprising one or more polymorphisms selected from the group consisting of thymine at PS1, thymine at PS5 and guanine at PS6.
7. An isolated polynucleotide comprising a nucleotide sequence which is a polymorphic variant of a reference sequence for the HTR1A cDNA or a fragment thereof, wherein the reference sequence comprises SEQ ID NO:4 and the polymorphic variant comprises at least one polymorphism selected from the group consisting of thymine at a position corresponding to nucleotide 423 and guanine at a position corresponding to nucleotide 464.
8. A recombinant organism transformed or transfected with the isolated polynucleotide of claim 7, wherein the organism expresses a 5-hydroxytryptamine receptor 1A gene (HTR1A) protein encoded by the polymorphic variant sequence.
9. The recombinant organism of claim 8 which is a nonhuman transgenic animal.
10. An isolated polypeptide comprising an amino acid sequence which is a polymorphic variant of a reference sequence for the HTR1A protein or a fragment thereof, wherein the reference sequence comprises SEQ ID NO:5 and the polymorphic variant comprises the glycine at a position corresponding to amino acid position 155.
11. An isolated antibody specific for and immunoreactive with the isolated polypeptide of claim 10.
12. A method for screening for drugs targeting the isolated polypeptide of claim 10 which comprises contacting the HTR1A polymorphic variant with a candidate agent and assaying for binding activity.
13. A composition comprising at least one genotyping oligonucleotide for detecting a polymorphism in the 5-hydroxytryptamine receptor 1A gene (HTR1A) gene at a polymorphic site selected from PS1,

PS5 and PS6.

14. The composition of claim 13, wherein the genotyping oligonucleotide is an allele-specific oligonucleotide that specifically hybridizes to an allele of the HTR1A gene at a region containing the polymorphic site.

15. The composition of claim 14, wherein the allele-specific oligonucleotide comprises a nucleotide sequence selected from the group consisting of of SEQ ID NOS:6-12, the complements of SEQ ID NOS: 6-11, and SEQ ID NOS:12-23.

16. The composition of claim 13, wherein the genotyping oligonucleotide is a primer-extension oligonucleotide.

17. A method for genotyping the 5-hydroxytryptamine receptor 1A gene (HTR1A) gene of an individual, comprising determining for the two copies of the HTR1A gene present in the individual the identity of the nucleotide pair at one or more polymorphic sites (PS) selected from PS1, PS5 and PS6.

18. The method of claim 17, wherein the determining step comprises:

(a) isolating from the individual a nucleic acid mixture comprising both copies of the HTR1A gene, or a fragment thereof, that are present in the individual;

(b) amplifying from the nucleic acid mixture a target region containing at least one of the polymorphic sites;

(c) hybridizing a primer extension oligonucleotide to one allele of the amplified target region;

(d) performing a nucleic acid template-dependent, primer extension reaction on the hybridized genotyping oligonucleotide in the presence of at least two different terminators of the reaction, wherein said terminators are complementary to the alternative nucleotides present at the polymorphic site; and

(e) detecting the presence and identity of the terminator in the extended genotyping oligonucleotide.

19. A method for haplotyping the 5-hydroxytryptamine receptor 1A gene (HTR1A) gene of an individual which comprises determining, for one copy of the HTR1A gene present in the individual, the identity of the nucleotide at one or more polymorphic sites (PS) selected from PS1, PS5 and PS6.

20. The method of claim 19, wherein the determining step comprises

(a) isolating from the individual a nucleic acid molecule containing only one of the two copies of the HTR1A gene, or a fragment thereof, that is present in the individual;

(b) amplifying from the nucleic acid molecule a target region containing at least one of the polymorphic sites;

(c) hybridizing a primer extension oligonucleotide to one allele of the amplified target region;

(d) performing a nucleic acid template-dependent, primer extension reaction on the hybridized genotyping oligonucleotide in the presence of at least two different terminators of the reaction,

wherein said terminators are complementary to the alternative nucleotides present at the polymorphic site; and

(e) detecting the presence and identity of the terminator in the extended genotyping oligonucleotide.

21. A method for predicting a haplotype pair for the 5-hydroxytryptamine receptor 1A gene (HTR1A) gene of an individual comprising:

(a) identifying an HTR1A genotype for the individual at two or more of polymorphic sites selected from PS1, PS5 and PS6;

(b) enumerating all possible haplotype pairs which are consistent with the genotype;

(c) accessing data containing the HTR1A haplotype pairs determined in a reference population; and

(d) assigning a haplotype pair to the individual that is consistent with the data.

22. A method for identifying an association between a trait and at least one genotype or haplotype of the 5-hydroxytryptamine receptor 1A gene gene which comprises comparing the frequency of the genotype or haplotype in a population exhibiting the trait with the frequency of the genotype or haplotype in a reference population, wherein the genotype or haplotype comprises a nucleotide pair or nucleotide located at one or more polymorphic sites selected from PS1, PS5 and PS6, wherein a higher frequency of the genotype or haplotype in the trait population than in the reference population indicates the trait is associated with the genotype or haplotype.

23 The method of claim 22, wherein the haplotype is selected from haplotype numbers 1-10 shown in Table 5.

24. The method of claim 23, wherein the trait is a clinical response to a drug targeting HTR1A .

25. A computer system for storing and analyzing polymorphism data for the 5-hydroxytryptamine receptor 1A gene gene, comprising:

(a) a central processing unit (CPU);

(b) a communication interface;

(c) a display device;

(d) an input device; and

(e) a database containing the polymorphism data;

wherein the polymorphism data comprises the genotypes and haplotype pairs shown in Table 4 and the haplotypes shown in Table 5.

26. A genome anthology for the 5-hydroxytryptamine receptor 1A gene (HTR1A) gene which comprises HTR1A isogenes defined by haplotypes 1-10 shown in Table 5.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 February 2001 (15.02.2001)

PCT

(10) International Publication Number
WO 01/10884 A1

(51) International Patent Classification?: C07H 21/04, C07K 14/00, 16/00, C12N 15/00

(21) International Application Number: PCT/US00/40519

(22) International Filing Date: 1 August 2000 (01.08.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/147,711 6 August 1999 (06.08.1999) US

(71) Applicant *(for all designated States except US)*: GENAISSANCE PHARMACEUTICALS, INC. [US/US]; Five Science Park, New Haven, CT 06511 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: DENTON, R., Rex [US/US]; 129 Hunters Trail, Madison, CT 06443 (US). KLIEM, Stefanie, E. [DE/US]; 1298 Hartford Turnpike, North Haven, CT 06473 (US). NANDABALAN, Krishnan [IN/US]; 228 Village Pond Road, Guilford, CT 06437 (US). STEPHENS, Joel, Claiborne [US/US]; 46 Crabapple Lane, Guilford, CT 06437 (US).

(74) Agent: HENDERSON, Melodie, W.; Genaissance Pharmaceuticals, Inc., Five Science Park, New Haven, CT 06511 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *With international search report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/10884 A1

(54) Title: DRUG TARGET ISOGENES: POLYMORPHISMS IN THE 5-HYDROXYTRYPTAMINE RECEPTOR 1A GENE

(57) Abstract: Polynucleotides comprising one or more of 3 novel single nucleotide polymorphisms in the human 5-hydroxytryptamine receptor 1A gene (HTR1A) gene are described. Compositions and methods for detecting one or more of these polymorphisms are also disclosed. In addition, various genotypes and haplotypes for HTR1A gene that exist in the population are described.

POLYMORPHISMS IN THE HTR1A GENE
(Accession No. Z11168)

```
ATTCTCCCTG AGGGAGTAAG GCTGGACTGT 12AGATGATAG CGGAGGTACC
GTTTTGTTGT TGTTGTCGTC GTTGTTCGTT TGTTTTTGGA GACGGAGTCT      100
CGCTCTGTCG CCCAGGCTGG AGTGCAATGG CGCGAGAACG GAGGTAGCTT
TTTAAAAACG AAGACACACT CGGTCTTCTT CCATCAATTA GCAATAATTG      200
GGAGACTGAC CCAGGACTGT TCACCTTCCC ATTCAGGCTC CCTATGCTTC
CTTTTCTCAT CTCCTATTGC CACTCTGGGA TGCTGACACG ATTTAAGAAT      300
TTGGCAGATA ATATGACCCA AGGAGTAGTT GGAATTCCCT CCCCCAAGTT
TTTCCAACCC CAGTTTTGCT GGGTTGGAGG CGGAGTTTAT TTGTTACAAC      400
CTTGGTCTGA CCGGCAGGGA CCTGGTGTGT GTAAGTGAGT TCTGAGTCTC
TGTTGACAAA AAGAGACTCG AATGCAAAGA CGCTGAGCTA GAGGGAGAGG      500
AGGGCGGGGA CCCAGAGGAA AGAGGCACTC CTCGGGGGTTG GGGAAGTATT
AGGAGGGGAG GGTTAGAGTG GGAGGGAAGG AGCTCGCTTT CGAAGCGACT      600
CACAGAGGGA TAAATAAAGG GAAGTGAGGA GGAAGAGGGA GACTTAAAGG
GAAGGCAGGT GGGGAGAAGG GGGACGAAAG AGGCAGAAGA GAGAGAAGAG      700
AGGAGGAGAG AGGGGGAGAG AGGGAAGGAA GGAAATAGCG AGAGGAGGGT
CACAGAGTGA CCGTGGAGGA TGGGGCTTCT CGGTTCTAGA TATTTCTGTG      800
ATTGGAGACT GTTTGCTAGT GGGGAGACTC CAGCTCCGGC AGCCAGTTCG
GGAGCGGCAA AGTAAAATGG ACAGCGACAG ACAGACGTTC CAGCCACCTC      900
TCCGCCGCCG GGAGATCCTG GAGCTGCTTT CAGGCCAACT CCAGTTTCCC
AGCTGGAGCT TCTGAACGCG CTGGACTGCG AGAGCCCAGG GAGCGCCTGA     1000
                                                    T
AAGCTGCTCC TCGGAGATAC CCTTCGCCGA AGCAGTAAGA ACTTCCTGCT
TGGGTCTCTG CATTCCCTTC CTCCGAAACT TCCCAGGAGA AGGGCGGAAG     1100
ACCCCAGGGG AAGGGGCGAG GCGAATCTTC GCGCTGCTTT TTCTTCCCTC
CCCCTTCCCG CGCCGGGCGC GCAGGCATGG ATGTGCTCAG CCCTGGTCAG     1200
GGAC                                                       1204
```

FIGURE 1

POLYMORPHISMS IN THE HTR1A GENE
(Accession No. M83181)

```
TCTAGATATT TCTGGGATTG GAGACTGTTT GCTAGTGGGG AGACTCCAGC
TCCGGCAGCC AGTTCGGGAG CGGCAAAGTA AAATGGACAG CGACAGACAG      100
ACGTTCCAGC CACCTCTCCG CCGCCGGGAG ATCCTGGAGC TGCTTTCAGG
CCAACTCCAG TTTCCCAGCT GGAGCTTCTG AACGCGCTGG ACTGCGAGAG      200
CCAGGGAGCG CCTGAAAGCT GCTCCTCGGA GATACCCTTC GCCGAAGCAG
TAAGAACTTC CTGCTTGGGT CTCTGCATTC CCTTCCTCCG AAACTTCCCA      300
GGAGAAGGGC GGAAGACCCC AGGGGAAGGG GCGAGGCGAA TCTTCGCGCT
GCTTTTTCTT CCCTCCCCCT TCCCGCGCCG GGCGCGCAGG CATGGATGTG      400
      [exon 1: 392..
CTCAGCCCTG GTCAGGGCAA CAACACCACA TCACCACCGG CTCCCTTTGA
                                           T
GACCGGCGGC AACACTACTG GTATCTCCGA CGTGACCGTC AGCTACCAAG      500
                G
TGATCACCTC TCTGCTGCTG GGCACGCTCA TCTTCTGCGC GGTGCTGGGC
AATGCGTGCG TGGTGGCTGC CATCGCCTTG GAGCGCTCCC TGCAGAACGT      600
GGCCAATTAT CTTATTGGCT CTTTGGCGGT CACCGACCTC ATGGTGTCGG
TGTTGGTGCT GCCCATGGCC GCGCTGTATC AGGTGCTCAA CAAGTGGACA      700
                                     A
CTGGGCCAGG TAACCTGCGA CCTGTTCATC GCCCTCGACG TGCTGTGCTG
CACCTCATCC ATCTTGCACC TGTGCGCCAT CGCGCTGGAC AGGTACTGGG      800
CCATCACGGA CCCCATCGAC TACGTGAACA AGAGGACGCC CCGGCGCGCC
          T
GCTGCGCTCA TCTCGCTCAC TTGGCTTATT GGCTTCCTCA TCTCTATCCC      900
   G
GCCCATGCTG GGCTGGCGCA CCCCGGAAGA CCGCTCGGAC CCCGACGCAT
                                              T
GCACCATTAG CAAGGATCAT GGCTACACTA TCTATTCCAC CTTTGGAGCT     1000
TTCTACATCC CGCTGCTGCT CATGCTGGTT CTCTATGGGC GCATATTCCG
                                              T
AGCTGCGCGC TTCCGCATCC GCAAGACGGT CAAAAAGGTG GAGAAGACCG     1100
GAGCGGACAC CCGCCATGGA GCATCTCCCG CCCCGCAGCC CAAGAAGAGT
GTGAATGGAG AGTCGGGGAG CAGGAACTGG AGGCTGGGCG TGGAGAGCAA     1200
GGCTGGGGGT GCTCTGTGCG CCAATGGCGC GGTGAGGCAA GGTGACGATG
      A
GCGCCGCCCT GGAGGTGATC GAGGTGCACC GAGTGGGCAA CTCCAAAGAG     1300
CACTTGCCTC TGCCCAGCGA GGCTGGTCCT ACCCCTTGTG CCCCCGCCTC
TTTCGAGAGG AAAAATGAGC GCAACGCCGA GGCGAAGCGC AAGATGGCCC     1400
TGGCCCGAGA GAGGAAGACA GTGAAGACGC TGGGCATCAT CATGGGCACC
TTCATCCTCT GCTGGCTGCC CTTCTTCATC GTGGCTCTTG TTCTGCCCTT     1500
CTGCGAGAGC AGCTGCCACA TGCCCACCCT GTTGGGCGCC ATAATCAATT
GGCTGGGCTA CTCCAACTCT CTGCTTAACC CCGTCATTTA CGCATACTTC     1600
AACAAGGACT TTCAAAACGC GTTTAAGAAG ATCATTAAGT GTAAGTTCTG
CCGCCAGTGA TGACGGAGGA GTAGCCGGCC AGTCGAGGCT ACAGGATCCG     1700
              ..1660]
TCCCATTCAC TATGCTTCCC CCAACCCTAG GGAATCAACA CTTAAGATAA
TTCGCCACTT CTCCTCTTTC TCTCTGCTCC GCTCACGGCT TGCAGACCTG     1800
GTCCCCTCCC CACTTCCTGC TCCACGGCAG GGCCCTTTGT GCAAAGGAGA
CCCAGCGGAG GAGCGTTGAG AGCCCAGGAA ATTCAGAGAG TTTGTGAGAA     1900
GCGACATTGG CTCAGACTTC GCCTGTATCA TCAGTTTT                  1938
```

FIGURE 2

POLYMORPHISMS IN A REFERENCE SEQUENCE FOR THE HTR1A GENE

```
ATTCTCCCTG AGGGAGTAAG GCTGGACTGT TAGATGATAG CGGAGGTACC
GTTTTGTTGT TGTTGTCGTC GTTGTTCGTT TGTTTTTGGA GACGGAGTCT      100
CGCTCTGTCG CCCAGGCTGG AGTGCAATGG CGCGAGAACG GAGGTAGCTT
TTTAAAAACG AAGACACACT CGGTCTTCTT CCAATCAATTA GCAATAATTG      200
GGAGACTGAC CCAGGACTGT TCACCTTCCC ATTCAGGCTC CCTATGCTTC
CTTTTCTCAT CTCCTATTGC CACTCTGGGA TGCTGACACG ATTTAAGAAT      300
TTGGCAGATA ATATGACCCA AGGAGTAGTT GGAATTCCCT CCCCCAAGTT
TTTCCAACCC CAGTTTTGCT GGGTTGGAGG CGGAGTTTAT TTGTTACAAC      400
CTTGGTCTGA CCGGCAGGGA CCTGGTGTGT GTAAGTGAGT TCTGAGTCTC
TGTTGACAAA AAGAGACTCG AATGCAAAGA CGCTGAGCTA GAGGGAGAGG      500
AGGGCGGGGA CCCAGAGGAA AGAGGCACTC CTCGGGGTTG GGGAAGTATT
AGGAGGGGAG GGTTAGAGTG GGAGGGAAGG AGCTCGCTTT CGAAGCGACT      600
CACAGAGGGA TAAATAAAGG GAAGTGAGGA GGAAGAGGGA GACTTAAAGG
GAAGGCAGGT GGGGAGAAGG GGGACGAAAG AGGCAGAAGA GAGAGAAGAG      700
AGGAGGAGAG AGGGGGAGAG AGGGAAGGAA GGAAATAGCG AGAGGAGGGT
CACAGAGTGA CCGTGGAGGA TGGGGCTTCT CGGTTCTAGA TATTTCTGGG      800
ATTGGAGACT GTTTGCTAGT GGGGAGACTC CAGCTCCGGC AGCCAGTTCG
GGAGCGGCAA AGTAAAATGG ACAGCGACAG ACAGACGTTC CAGCCACCTC      900
TCCGCCGCCG GGAGATCCTG GAGCTGCTTT CAGGCCAACT CCAGTTTCCC
AGCTGGAGCT TCTGAACGCG CTGGACTGCG AGAGCCAGGG AGCGCCTGAA     1000
                                                      T
AGCTGCTCCT CGGAGATACC CTTCGCCGAA GCAGTAAGAA CTTCCTGCTT
GGGTCTCTGC ATTCCCTTCC TCCGAAACTT CCCAGGAGAA GGGCGGAAGA     1100
CCCCAGGGGA AGGGGCGAGG CGAATCTTCG CGCTGCTTTT TCTTCCCTCC
CCCTTCCCGC GCCGGGCGCG CAGGCATGGA TGTGCTCAGC CCTGGTCAGG     1200
          [exon 1:1176..
GCAACAACAC CACATCACCA CCGGCTCCCT TTGAGACCGG CGGCAACACT
                     T
ACTGGTATCT CCGACGTGAC CGTCAGCTAC CAAGTGATCA CCTCTCTGCT     1300
   G
GCTGGGCACG CTCATCTTCT GCGCGGTGCT GGGCAATGCG TGCGTGGTGG
CTGCCATCGC CTTGGAGCGC TCCCTGCAGA ACGTGGCCAA TTATCTTATT     1400
GGCTCTTTGG CGGTCACCGA CCTCATGGTG TCGGTGTTGG TGCTGCCCAT
GGCCGCGCTG TATCAGGTGC TCAACAAGTG GACACTGGGC CAGGTAACCT     1500
            A
GCGACCTGTT CATCGCCCTC GACGTGCTGT GCTGCACCTC ATCCATCTTG
CACCTGTGCG CCATCGCGCT GGACAGGTAC TGGGCCATCA CGGACCCCAT     1600
                                                      T
CGACTACGTG AACAAGAGGA CGCCCCGGCG CGCCGCTGCG CTCATCTCGC
                                    G
TCACTTGGCT TATTGGCTTC CTCATCTCTA TCCCGCCCAT GCTGGGCTGG     1700
CGCACCCCGG AAGACCGCTC GGACCCCGAC GCATGCACCA TTAGCAAGGA
                  T
TCATGGCTAC ACTATCTATT CCACCTTTGG AGCTTTCTAC ATCCCGCTGC     1800
TGCTCATGCT GGTTCTCTAT GGGCGCATAT TCCGAGCTGC GCGCTTCCGC
                           T
ATCCGCAAGA CGGTCAAAAA GGTGGAGAAG ACCGGAGCGG ACACCCGCCA     1900
TGGAGCATCT CCCGCCCCGC AGCCCAAGAA GAGTGTGAAT GGAGAGTCGG
GGAGCAGGAA CTGGAGGCTG GGCGTGGAGA GCAAGGCTGG GGGTGCTCTG     2000
                                             A
TGCGCCAATG GCGCGGTGAG GCAAGGTGAC GATGGCGCCG CCCTGGAGGT
GATCGAGGTG CACCGAGTGG GCAACTCCAA AGAGCACTTG CCTCTGCCCA     2100
GCGAGGCTGG TCCTACCCCT TGTGCCCCCG CCTCTTTCGA GAGGAAAAAT
GAGCGCAACG CCGAGGCGAA GCGCAAGATG GCCCTGGCCC GAGAGAGGAA     2200
GACAGTGAAG ACGCTGGGCA TCATCATGGG CACCTTCATC CTCTGCTGGC
```

FIGURE 3A

```
TGCCCTTCTT CATCGTGGCT CTTGTTCTGC CCTTCTGCGA GAGCAGCTGC      2300
CACATGCCCA CCCTGTTGGG CGCCATAATC AATTGGCTGG GCTACTCCAA
CTCTCTGCTT AACCCCGTCA TTTACGCATA CTTCAACAAG GACTTTCAAA      2400
ACGCGTTTAA GAAGATCATT AAGTGTAAGT TCTGCCGCCA GTGATGACGG
                                                  2444]
AGGAGTAGCC GGCCAGTCGA GGCTACAGGA TCCGTCCCAT TCACTATGCT      2500
TCCCCCAACC CTAGGGAATC AACACTTAAG ATAATTCGCC ACTTCTCCTC
TTTCTCTCTG CTCCGCTCAC GGCTTGCAGA CCTGGTCCCC TCCCCACTTC      2600
CTGCTCCACG GCAGGGCCCT TTGTGCAAAG GAGACCCAGC GGAGGAGCGT
TGAGAGCCCA GGAAATTCAG AGAGTTTGTG AGAAGCGACA TTGGCTCAGA      2700
CTTCGCCTGT ATCATCAGTT TT                                   2722
```

FIGURE 3B

POLYMORPHISMS IN THE CODING SEQUENCE OF HTR1A

```
ATGGATGTGC TCAGCCCTGG TCAGGGCAAC AACACCACAT CACCACCGGC
                                                    T
TCCCTTTGAG ACCGGCGGCA ACACTACTGG TATCTCCGAC GTGACCGTCA      100
                                    G
GCTACCAAGT GATCACCTCT CTGCTGCTGG GCACGCTCAT CTTCTGCGCG
GTGCTGGGCA ATGCGTGCGT GGTGGCTGCC ATCGCCTTGG AGCGCTCCCT      200
GCAGAACGTG GCCAATTATC TTATTGGCTC TTTGGCGGTC ACCGACCTCA
TGGTGTCGGT GTTGGTGCTG CCCATGGCCG CGCTGTATCA GGTGCTCAAC      300
                                                    A
AAGTGGACAC TGGGCCAGGT AACCTGCGAC CTGTTCATCG CCCTCGACGT
GCTGTGCTGC ACCTCATCCA TCTTGCACCT GTGCGCCATC GCGCTGGACA      400
GGTACTGGGC CATCACGGAC CCCATCGACT ACGTGAACAA GAGGACGCCC
                      T
CGGCGCGCCG CTGCGCTCAT CTCGCTCACT TGGCTTATTG GCTTCCTCAT      500
           G
CTCTATCCCG CCCATGCTGG GCTGGCGCAC CCCGGAAGAC CGCTCGGACC
CCGACGCATG CACCATTAGC AAGGATCATG GCTACACTAT CTATTCCACC      600
 T
TTTGGAGCTT TCTACATCCC GCTGCTGCTC ATGCTGGTTC TCTATGGGCG
CATATTCCGA GCTGCGCGCT TCCGCATCCG CAAGACGGTC AAAAAGGTGG      700
      T
AGAAGACCGG AGCGGACACC CGCCATGGAG CATCTCCCGC CCCGCAGCCC
AAGAAGAGTG TGAATGGAGA GTCGGGGAGC AGGAACTGGA GGCTGGGCGT      800
GGAGAGCAAG GCTGGGGGTG CTCTGTGCGC CAATGGCGCG GTGAGGCAAG
                  A
GTGACGATGG CGCCGCCCTG GAGGTGATCG AGGTGCACCG AGTGGGCAAC      900
TCCAAAGAGC ACTTGCCTCT GCCCAGCGAG GCTGGTCCTA CCCCTTGTGC
CCCCGCCTCT TTCGAGAGGA AAAATGAGCG CAACGCCGAG GCGAAGCGCA     1000
AGATGGCCCT GGCCCGAGAG AGGAAGACAG TGAAGACGCT GGGCATCATC
ATGGGCACCT TCATCCTCTG CTGGCTGCCC TTCTTCATCG TGGCTCTTGT     1100
TCTGCCCTTC TGCGAGAGCA GCTGCCACAT GCCCACCCTG TTGGGCGCCA
TAATCAATTG GCTGGGCTAC TCCAACTCTC TGCTTAACCC CGTCATTTAC     1200
GCATACTTCA ACAAGGACTT TCAAAACGCG TTTAAGAAGA TCATTAAGTG
TAAGTTCTGC CGCCAGTGA                                      1269
```

FIGURE 4

ISOFORMS OF THE HTR1A PROTEIN

```
MDVLSPGQGN NTTSPPAPFE TGGNTTGISD VTVSYQVITS LLLGTLIFCA
             L           V
VLGNACVVAA IALERSLQNV ANYLIGSLAV TDLMVSVLVL PMAALYQVLN      100
KWTLGQVTCD LFIALDVLCC TSSILHLCAI ALDRYWAITD PIDYVNKRTP
RRAAALISLT WLIGFLISIP PMLGWRTPED RSDPDACTIS KDHGYTIYST      200
   G
FGAFYIPLLL MLVLYGRIFR AARFRIRKTV KKVEKTGADT RHGASPAPQP
                 L
KKSVNGESGS RNWRLGVESK AGGALCANGA VRQGDDGAAL EVIEVHRVGN      300
                         D
SKEHLPLPSE AGPTPCAPAS FERKNERNAE AKRKMALARE RKTVKTLGII
MGTFILCWLP FFIVALVLPF CESSCHMPTL LGAIINWLGY SNSLLNPVIY      400
AYFNKDFQNA FKKIIKCKFC RQ                                    422
```

FIGURE 5

Practitioner's Docket No. MWH-0006PCT *PATENT*

COMBINED DECLARATION AND POWER OF ATTORNEY

(ORIGINAL, DESIGN, NATIONAL STAGE OF PCT, SUPPLEMENTAL, DIVISIONAL, CONTINUATION, OR C-I-P)

As a below named inventor, I hereby declare that:

TYPE OF DECLARATION

This declaration is for a national stage of PCT application.

INVENTORSHIP IDENTIFICATION

My residence, post office address and citizenship are as stated below, beneath my name. I believe that I am an original, first and joint inventor of the subject matter that is claimed, and for which a patent is sought on the invention entitled:

TITLE OF INVENTION

Drug Target Isogenes: Polymorphisms in the 5-Hydroxytryptamine Receptor 1A Gene

SPECIFICATION IDENTIFICATION

The specification was described and claimed in PCT International Application No. PCT/US00/40519 filed on August 1, 2000 and was amended under PCT Article 34(2)(b) on March 5, 2001.

ACKNOWLEDGMENT OF REVIEW OF PAPERS AND DUTY OF CANDOR

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by the amendment referred to above.

I acknowledge the duty to disclose information, which is material to patentability as defined in 37, Code of Federal Regulations, Section 1.56, and which is material to the examination of this application, namely, information where there is a substantial likelihood that a reasonable Examiner would consider it important in deciding whether to allow the application to issue as a patent, and in compliance with this duty, there is attached an information disclosure statement, in accordance with 37 C.F.R. Section 1.98.

CLAIM FOR BENEFIT OF PRIOR U.S. PROVISIONAL APPLICATION(S)
(35 U.S.C. Section 119(e))

I hereby claim the benefit under Title 35, United States Code, Section 119(e) of any United States provisional application(s) listed below:

PROVISIONAL APPLICATION NUMBER	FILING DATE
60/147,711	August 6, 1999

POWER OF ATTORNEY

I hereby appoint the following practitioner(s) to prosecute this application and transact all business in the Patent and Trademark Office connected therewith.

APPOINTED PRACTITIONER(S)	REGISTRATION NUMBER(S)
Gisela M. Field	47,562
Melodie W. Henderson	~~37,848~~ 4
Nathaniel D. Judisch	50,143
Sandra L. Shaner	47,934

I hereby appoint the practitioner(s) associated with the Customer Number provided below to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith.

SEND CORRESPONDENCE TO

Gisela M. Field
5 Science Park
New Haven, CT 06511
USA

Customer Number 25106

DIRECT TELEPHONE CALLS TO:

Gisela M. Field
203-786-3473

DECLARATION

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

SIGNATURE(S)

1-00

R. Rex Denton
Inventor's signature R. Rex Denton
Date 2-1-02 **Country of Citizenship** USA
Residence Madison, CT
Post Office Address 129 Hunter's Trail, Madison, CT 06443 USA CT

2-00

Stefanie E. Kliem
Inventor's signature Stefanie Kliem
Date 25 January 2002 **Country of Citizenship** Germany
Residence Oberusel Germany
Post Office Address Kiefernweg 37, Oberusel 61440 Germany DE

3-00

Krishnan Nandabalan
Inventor's signature K. Nandabalan
Date 2/4/2 **Country of Citizenship** USA
Residence Guilford, CT
Post Office Address 228 Village Pond Road, Guilford, CT 06437 USA CT

4-00

J. Claiborne Stephens
Inventor's signature J. Claiborne Stephens CT
Date 2-1-02 **Country of Citizenship** USA
Residence Guilford, CT
Post Office Address 46 Crabapple Lane, Guilford, CT 06437 USA

SEQUENCE LISTING

<110> Denten, R. Rex
Nandabalan, Krishnan
Kleim, Stepfanie
Stephens, J. Claiborne

<120> DRUG TARGET ISOGENES: POLYMORPHISMS IN THE 5-HYDROXYTRYPTAMINE RECEPTOR 1A GENE

<130> MWH-0006PCT HTR1A

<140> tba
<141> 2000-08-01

<150> 60/147,711
<151> 1999-08-06

<160> 53

<170> PatentIn Ver. 2.1

<210> 1
<211> 1204
<212> DNA
<213> Homo sapiens

<400> 1
attctccctg agggagtaag gctggactgt tagatgatag cggaggtacc gttttgttgt 60
tgttgtcgtc gttgttcgtt tgtttttgga gacggagtct cgctctgtcg cccaggctgg 120
agtgcaatgg cgcgagaacg gaggtagctt tttaaaaacg aagacacact cggtcttctt 180
ccatcaatta gcaatatttg ggagactgac ccaggactgt tcaccttccc attcaggctc 240
cctatgcttc cttttctcat ctcctattgc cactctggga tgctgacacg atttaagaat 300
ttggcagata atatgaccca aggagtagtt ggaattccct cccccaagtt tttccaaccc 360
cagttttgct gggttggagg cggagtttat ttgttacaac cttggtctga ccggcaggga 420
cctggtgtgt gtaagtgagt tctgagtctc tgttgacaaa aagagactcg aatgcaaaga 480
cgctgagcta gagggagagg agggcgggga cccagaggaa agaggcactc ctcggggttg 540
gggaagtatt aggaggggag ggttagagtg ggagggaagg agctcgcttt cgaagcgact 600
cacagaggga taaataaagg gaagtgagga ggaagaggga gacttaaagg gaaggcaggt 660
ggggagaagg gggacgaaag aggcagaaga gagagaagag aggaggagag agggggagag 720
agggaaggaa ggaaatagcg agaggagggt cacagagtga ccgtggagga tggggcttct 780
cggttctaga tatttctgtg attggagact gtttgctagt ggggagactc cagctccggc 840
agccagttcg ggagcggcaa agtaaaatgg acagcgacag acagacgttc cagccacctc 900
tccgccgccg ggagatcctg gagctgcttt caggccaact ccagtttccc agctggagct 960
tctgaacgcg ctggactgcg agagcccagg gagcgcctga aagctgctcc tcggagatac 1020
ccttcgccga agcagtaaga acttcctgct tgggtctctg cattcccttc ctccgaaact 1080
tcccaggaga agggcggaag accccagggg aaggggcgag gcgaatcttc gcgctgcttt 1140
ttcttccctc ccccttcccg cgccgggcgc gcaggcatgg atgtgctcag ccctggtcag 1200

ggac 1204

<210> 2
<211> 1938
<212> DNA
<213> Homo sapiens

<400> 2
tctagatatt tctgggattg gagactgttt gctagtgggg agactccagc tccggcagcc 60
agttcgggag cggcaaagta aaatggacag cgacagacag acgttccagc cacctctccg 120
ccgccgggag atcctggagc tgctttcagg ccaactccag tttcccagct ggagcttctg 180
aacgcgctgg actgcgagag ccagggagcg cctgaaagct gctcctcgga gatacccttc 240
gccgaagcag taagaacttc ctgcttgggt ctctgcattc ccttcctccg aaacttccca 300
ggagaagggc ggaagacccc aggggaaggg gcgaggcgaa tcttcgcgct gctttttctt 360
ccctccccct tcccgcgccg ggcgcgcagg catggatgtg ctcagccctg gtcagggcaa 420
caacaccaca tcaccaccgg ctccctttga gaccggcggc aacactactg gtatctccga 480
cgtgaccgtc agctaccaag tgatcacctc tctgctgctg ggcacgctca tcttctgcgc 540
ggtgctgggc aatgcgtgcg tggtggctgc catcgccttg gagcgctccc tgcagaacgt 600
ggccaattat cttattggct ctttggcggt caccgacctc atggtgtcgg tgttggtgct 660
gcccatggcc gcgctgtatc aggtgctcaa caagtggaca ctgggccagg taacctgcga 720
cctgttcatc gccctcgacg tgctgtgctg cacctcatcc atcttgcacc tgtgcgccat 780
cgcgctggac aggtactggg ccatcacgga ccccatcgac tacgtgaaca agaggacgcc 840
ccggcgcgcc gctgcgctca tctcgctcac ttggcttatt ggcttcctca tctctatccc 900
gcccatgctg ggctggcgca ccccggaaga ccgctcggac cccgacgcat gcaccattag 960
caaggatcat ggctacacta tctattccac ctttggagct ttctacatcc cgctgctgct 1020
catgctggtt ctctatgggc gcatattccg agctgcgcgc ttccgcatcc gcaagacggt 1080
caaaaaggtg gagaagaccg gagcggacac ccgccatgga gcatctcccg ccccgcagcc 1140
caagaagagt gtgaatggag agtcggggag caggaactgg aggctgggcg tggagagcaa 1200
ggctgggggt gctctgtgcg ccaatggcgc ggtgaggcaa ggtgacgatg gcgccgccct 1260
ggaggtgatc gaggtgcacc gagtgggcaa ctccaaagag cacttgcctc tgcccagcga 1320
ggctggtcct acccctgtg cccccgcctc tttcgagagg aaaaatgagc gcaacgccga 1380
ggcgaagcgc aagatggccc tggcccgaga gaggaagaca gtgaagacgc tgggcatcat 1440
catgggcacc ttcatcctct gctggctgcc cttcttcatc gtggctcttg ttctgccctt 1500
ctgcgagagc agctgccaca tgcccaccct gttgggcgcc ataatcaatt ggctgggcta 1560
ctccaactct ctgcttaacc ccgtcattta cgcatacttc aacaaggact ttcaaaacgc 1620
gtttaagaag atcattaagt gtaagttctg ccgccagtga tgacggagga gtagccggcc 1680
agtcgaggct acaggatccg tcccattcac tatgcttccc ccaaccctag ggaatcaaca 1740
cttaagataa ttcgccactt ctcctctttc tctctgctcc gctcacggct tgcagaccctg 1800
gtcccctccc cacttcctgc tccacggcag ggccctttgt gcaaaggaga cccagcggag 1860
gagcgttgag agcccaggaa attcagagag tttgtgagaa gcgacattgg ctcagacttc 1920
gcctgtatca tcagtttt 1938

<210> 3
<211> 2722
<212> DNA
<213> Homo sapiens

<400> 3
attctccctg agggagtaag gctggactgt tagatgatag cggaggtacc gttttgttgt 60
tgttgtcgtc gttgttcgtt tgtttttgga gacggagtct cgctctgtcg cccaggctgg 120
agtgcaatgg cgcgagaacg gaggtagctt tttaaaaacg aagacacact cggtcttctt 180
ccatcaatta gcaataattg ggagactgac ccaggactgt tcaccttccc attcaggctc 240
cctatgcttc cttttctcat ctcctattgc cactctggga tgctgacacg atttaagaat 300
ttggcagata atatgaccca aggagtagtt ggaattccct cccccaagtt tttccaaccc 360
cagttttgct gggttggagg cggagtttat ttgttacaac cttggtctga ccggcaggga 420
cctggtgtgt gtaagtgagt tctgagtctc tgttgacaaa aagagactcg aatgcaaaga 480
cgctgagcta gagggagagg agggcgggga cccagaggaa agaggcactc ctcggggttg 540
gggaagtatt aggaggggag ggttagagtg ggagggaagg agctcgcttt cgaagcgact 600
cacagaggga taaataaagg gaagtgagga ggaagaggga gacttaaagg gaaggcaggt 660
ggggagaagg gggacgaaag aggcagaaga gagagaagag aggaggagag aggggagag 720
agggaaggaa ggaaatagcg agaggagggt cacagagtga ccgtggagga tggggcttct 780
cggttctaga tatttctggg attggagact gtttgctagt ggggagactc cagctccggc 840
agccagttcg ggagcggcaa agtaaaatgg acagcgacag acagacgttc cagccacctc 900
tccgccgccg ggagatcctg gagctgcttt caggccaact ccagtttccc agctggagct 960
tctgaacgcg ctggactgcg agagccaggg agcgcctgaa agctgctcct cggagatacc 1020
cttcgccgaa gcagtaagaa cttcctgctt gggtctctgc attcccttcc tccgaaactt 1080
cccaggagaa gggcggaaga ccccagggga aggggcgagg cgaatcttcg cgctgctttt 1140
tcttccctcc ccctttccgc gccgggcgcg caggcatgga tgtgctcagc cctggtcagg 1200
gcaacaacac cacatcacca ccggctccct ttgagaccgg cggcaacact actggtatct 1260
ccgacgtgac cgtcagctac caagtgatca cctctctgct gctgggcacg ctcatcttct 1320
gcgcggtgct gggcaatgcg tgcgtggtgg ctgccatcgc cttggagcgc tccctgcaga 1380
acgtggccaa ttatcttatt ggctctttgg cggtcaccga cctcatggtg tcggtgttgg 1440
tgctgcccat ggccgcgctg tatcaggtgc tcaacaagtg gacactgggc caggtaacct 1500
gcgacctgtt catcgccctc gacgtgctgt gctgcacctc atccatcttg cacctgtgcg 1560
ccatcgcgct ggacaggtac tgggccatca cggaccccat cgactacgtg aacaagagga 1620
cgccccggcg cgccgctgcg ctcatctcgc tcacttggct tattggcttc ctcatctcta 1680
tcccgcccat gctgggctgg cgcaccccgg aagaccgctc ggaccccgac gcatgcacca 1740
ttagcaagga tcatggctac actatctatt ccacctttgg agctttctac atcccgctgc 1800
tgctcatgct ggttctctat gggcgcatat tccgagctgc gcgcttccgc atccgcaaga 1860
cggtcaaaaa ggtggagaag accggagcgg acacccgcca tggagcatct cccgccccgc 1920
agcccaagaa gagtgtgaat ggagagtcgg ggagcaggaa ctggaggctg ggcgtggaga 1980
gcaaggctgg gggtgctctg tgcgccaatg gcgcggtgag gcaaggtgac gatggcgccg 2040
ccctggaggt gatcgaggtg caccgagtgg gcaactccaa agagcacttg cctctgccca 2100
gcgaggctgg tcctacccct tgtgcccccg cctctttcga gaggaaaaat gagcgcaacg 2160
ccgaggcgaa gcgcaagatg gccctggccc gagagaggaa gacagtgaag acgctgggca 2220
tcatcatggg caccttcatc ctctgctggc tgcccttctt catcgtggct cttgttctgc 2280
ccttctgcga gagcagctgc cacatgccca ccctgttggg cgccataatc aattggctgg 2340
gctactccaa ctctctgctt aaccccgtca tttacgcata cttcaacaag gactttcaaa 2400
acgcgtttaa gaagatcatt aagtgtaagt tctgccgcca gtgatgacgg aggagtagcc 2460
ggccagtcga ggctacagga tccgtcccat tcactatgct tcccccaacc ctagggaatc 2520
aacacttaag ataattcgcc acttctcctc tttctctctg ctccgctcac ggcttgcaga 2580
cctggtcccc tccccacttc ctgctccacg gcagggccct ttgtgcaaag gagacccagc 2640
ggaggagcgt tgagagccca ggaaattcag agagtttgtg agaagcgaca ttggctcaga 2700
cttcgcctgt atcatcagtt tt 2722

<210> 4
<211> 1269
<212> DNA
<213> Homo sapiens

<400> 4
atggatgtgc tcagccctgg tcagggcaac aacaccacat caccaccggc tccctttgag 60
accggcggca acactactgg tatctccgac gtgaccgtca gctaccaagt gatcacctct 120
ctgctgctgg gcacgctcat cttctgcgcg gtgctgggca atgcgtgcgt ggtggctgcc 180
atcgccttgg agcgctccct gcagaacgtg gccaattatc ttattggctc tttggcggtc 240
accgacctca tggtgtcggt gttggtgctg cccatggccg cgctgtatca ggtgctcaac 300
aagtggacac tgggccaggt aacctgcgac ctgttcatcg ccctcgacgt gctgtgctgc 360
acctcatcca tcttgcacct gtgcgccatc gcgctggaca ggtactgggc catcacggac 420
cccatcgact acgtgaacaa gaggacgccc cggcgcgccg ctgcgctcat ctcgctcact 480
tggcttattg gcttcctcat ctctatcccg cccatgctgg gctggcgcac cccggaagac 540
cgctcggacc ccgacgcatg caccattagc aaggatcatg gctacactat ctattccacc 600
tttggagctt tctacatccc gctgctgctc atgctggttc tctatgggcg catattccga 660
gctgcgcgct tccgcatccg caagacggtc aaaaaggtgg agaagaccgg agcggacacc 720
cgccatggag catctcccgc cccgcagccc aagaagagtg tgaatggaga gtcggggagc 780
aggaactgga ggctgggcgt ggagagcaag gctgggggtg ctctgtgcgc caatggcgcg 840
gtgaggcaag gtgacgatgg cgccgccctg gaggtgatcg aggtgcaccg agtgggcaac 900
tccaaagagc acttgcctct gcccagcgag gctggtccta ccccttgtgc ccccgcctct 960
ttcgagagga aaaatgagcg caacgccgag gcgaagcgca agatggccct ggcccgagag 1020
aggaagacag tgaagacgct gggcatcatc atgggcacct tcatcctctg ctggctgccc 1080
ttcttcatcg tggctcttgt tctgcccttc tgcgagagca gctgccacat gcccaccctg 1140
ttgggcgcca taatcaattg gctgggctac tccaactctc tgcttaaccc cgtcatttac 1200
gcatacttca acaaggactt tcaaaacgcg tttaagaaga tcattaagtg taagttctgc 1260
cgccagtga 1269

<210> 5
<211> 422
<212> PRT
<213> Homo sapiens

<400> 5
Met Asp Val Leu Ser Pro Gly Gln Gly Asn Asn Thr Thr Ser Pro Pro
1 5 10 15

Ala Pro Phe Glu Thr Gly Gly Asn Thr Thr Gly Ile Ser Asp Val Thr
20 25 30

Val Ser Tyr Gln Val Ile Thr Ser Leu Leu Leu Gly Thr Leu Ile Phe
35 40 45

Cys Ala Val Leu Gly Asn Ala Cys Val Val Ala Ala Ile Ala Leu Glu

```
     50                  55                  60

Arg Ser Leu Gln Asn Val Ala Asn Tyr Leu Ile Gly Ser Leu Ala Val
 65                  70                  75                  80

Thr Asp Leu Met Val Ser Val Leu Val Leu Pro Met Ala Ala Leu Tyr
                 85                  90                  95

Gln Val Leu Asn Lys Trp Thr Leu Gly Gln Val Thr Cys Asp Leu Phe
            100                 105                 110

Ile Ala Leu Asp Val Leu Cys Cys Thr Ser Ser Ile Leu His Leu Cys
        115                 120                 125

Ala Ile Ala Leu Asp Arg Tyr Trp Ala Ile Thr Asp Pro Ile Asp Tyr
    130                 135                 140

Val Asn Lys Arg Thr Pro Arg Arg Ala Ala Ala Leu Ile Ser Leu Thr
145                 150                 155                 160

Trp Leu Ile Gly Phe Leu Ile Ser Ile Pro Pro Met Leu Gly Trp Arg
                165                 170                 175

Thr Pro Glu Asp Arg Ser Asp Pro Asp Ala Cys Thr Ile Ser Lys Asp
            180                 185                 190

His Gly Tyr Thr Ile Tyr Ser Thr Phe Gly Ala Phe Tyr Ile Pro Leu
        195                 200                 205

Leu Leu Met Leu Val Leu Tyr Gly Arg Ile Phe Arg Ala Ala Arg Phe
    210                 215                 220

Arg Ile Arg Lys Thr Val Lys Lys Val Glu Lys Thr Gly Ala Asp Thr
225                 230                 235                 240

Arg His Gly Ala Ser Pro Ala Pro Gln Pro Lys Lys Ser Val Asn Gly
                245                 250                 255

Glu Ser Gly Ser Arg Asn Trp Arg Leu Gly Val Glu Ser Lys Ala Gly
            260                 265                 270

Gly Ala Leu Cys Ala Asn Gly Ala Val Arg Gln Gly Asp Asp Gly Ala
        275                 280                 285

Ala Leu Glu Val Ile Glu Val His Arg Val Gly Asn Ser Lys Glu His
    290                 295                 300

Leu Pro Leu Pro Ser Glu Ala Gly Pro Thr Pro Cys Ala Pro Ala Ser
```

305 310 315 320

Phe Glu Arg Lys Asn Glu Arg Asn Ala Glu Ala Lys Arg Lys Met Ala
325 330 335

Leu Ala Arg Glu Arg Lys Thr Val Lys Thr Leu Gly Ile Ile Met Gly
340 345 350

Thr Phe Ile Leu Cys Trp Leu Pro Phe Phe Ile Val Ala Leu Val Leu
355 360 365

Pro Phe Cys Glu Ser Ser Cys His Met Pro Thr Leu Leu Gly Ala Ile
370 375 380

Ile Asn Trp Leu Gly Tyr Ser Asn Ser Leu Leu Asn Pro Val Ile Tyr
385 390 395 400

Ala Tyr Phe Asn Lys Asp Phe Gln Asn Ala Phe Lys Lys Ile Ile Lys
405 410 415

Cys Lys Phe Cys Arg Gln
420

<210> 6
<211> 15
<212> DNA
<213> Homo sapiens

<400> 6
ggagcgcctg aaagc 15

<210> 7
<211> 15
<212> DNA
<213> Homo sapiens

<400> 7
ggagcgcttg aaagc 15

<210> 8
<211> 15
<212> DNA
<213> Homo sapiens

<400> 8

cggaccccat cgact 15

<210> 9
<211> 15
<212> DNA
<213> Homo sapiens

<400> 9
cggaccctat cgact 15

<210> 10
<211> 15
<212> DNA
<213> Homo sapiens

<400> 10
gccgctgcgc tcatc 15

<210> 11
<211> 15
<212> DNA
<213> Homo sapiens

<400> 11
gccgctgggc tcatc 15

<210> 12
<211> 15
<212> DNA
<213> Homo sapiens

<400> 12
gcccagggag cgcct 15

<210> 13
<211> 15
<212> DNA
<213> Homo sapiens

<400> 13
ggagcagctt tcagg 15

<210> 14
<211> 15
<212> DNA
<213> Homo sapiens

<400> 14
gcccagggag cgctt 15

<210> 15
<211> 16
<212> DNA
<213> Homo sapiens

<400> 15
ggagcagctt tcaagc 16

<210> 16
<211> 15
<212> DNA
<213> Homo sapiens

<400> 16
ccatcacgga cccca 15

<210> 17
<211> 15
<212> DNA
<213> Homo sapiens

<400> 17
tcacgtagtc gatgg 15

<210> 18
<211> 15
<212> DNA
<213> Homo sapiens

<400> 18
ccatcacgga cccta 15

<210> 19
<211> 16
<212> DNA

<213> Homo sapiens

<400> 19
tcacgtagtc gatagg 16

<210> 20
<211> 15
<212> DNA
<213> Homo sapiens

<400> 20
cggcgcgccg ctgcg 15

<210> 21
<211> 15
<212> DNA
<213> Homo sapiens

<400> 21
gagcgagatg agcgc 15

<210> 22
<211> 15
<212> DNA
<213> Homo sapiens

<400> 22
cggcgcgccg ctggg 15

<210> 23
<211> 16
<212> DNA
<213> Homo sapiens

<400> 23
gagcgagatg agccca 16

<210> 24
<211> 10
<212> DNA
<213> Homo sapiens

<400> 24

cagggagcgc 10

<210> 25
<211> 10
<212> DNA
<213> Homo sapiens

<400> 25
gcagctttca 10

<210> 26
<211> 10
<212> DNA
<213> Homo sapiens

<400> 26
tcacggaccc 10

<210> 27
<211> 10
<212> DNA
<213> Homo sapiens

<400> 27
cgtagtcgat 10

<210> 28
<211> 10
<212> DNA
<213> Homo sapiens

<400> 28
cgcgccgctg 10

<210> 29
<211> 10
<212> DNA
<213> Homo sapiens

<400> 29
cgagatgagc 10

<210> 30
<211> 26
<212> DNA
<213> Homo sapiens

<400> 30
aagaggcaga agagagagaa gagagg 26

<210> 31
<211> 22
<212> DNA
<213> Homo sapiens

<400> 31
aagtttcgga ggaagggaat gc 22

<210> 32
<211> 22
<212> DNA
<213> Homo sapiens

<400> 32
aaagctgctc ctcggagata cc 22

<210> 33
<211> 22
<212> DNA
<213> Homo sapiens

<400> 33
cacaggtgca agatggatga gg 22

<210> 34
<211> 25
<212> DNA
<213> Homo sapiens

<400> 34
cactactggt atctccgacg tgacc 25

<210> 35
<211> 23
<212> DNA

<213> Homo sapiens

<400> 35
ggaatatgcg cccatagaga acc 23

<210> 36
<211> 21
<212> DNA
<213> Homo sapiens

<400> 36
gctgcacctc atccatcttg c 21

<210> 37
<211> 23
<212> DNA
<213> Homo sapiens

<400> 37
aggcaagtgc tctttggagt tgc 23

<210> 38
<211> 23
<212> DNA
<213> Homo sapiens

<400> 38
ctgctcatgc tggttctcta tgg 23

<210> 39
<211> 22
<212> DNA
<213> Homo sapiens

<400> 39
acggggttaa gcagagagtt gg 22

<210> 40
<211> 22
<212> DNA
<213> Homo sapiens

<400> 40

cgagtgggca actccaaaga gc 22

<210> 41
<211> 20
<212> DNA
<213> Homo sapiens

<400> 41
aggggaccag gtctgcaagc 20

<210> 42
<211> 20
<212> DNA
<213> Homo sapiens

<400> 42
gaggggggaga gagggaagga 20

<210> 43
<211> 20
<212> DNA
<213> Homo sapiens

<400> 43
tgcagagacc caagcaggaa 20

<210> 44
<211> 20
<212> DNA
<213> Homo sapiens

<400> 44
gggtctctgc attcccttcc 20

<210> 45
<211> 20
<212> DNA
<213> Homo sapiens

<400> 45
cgatgaacag gtcgcaggtt 20

<210> 46
<211> 20
<212> DNA
<213> Homo sapiens

<400> 46
ctgggcacgc tcatcttctg 20

<210> 47
<211> 20
<212> DNA
<213> Homo sapiens

<400> 47
atgagcagca gcgggatgta 20

<210> 48
<211> 20
<212> DNA
<213> Homo sapiens

<400> 48
acaggtactg ggccatcacg 20

<210> 49
<211> 20
<212> DNA
<213> Homo sapiens

<400> 49
gtgcacctcg atcacctcca 20

<210> 50
<211> 18
<212> DNA
<213> Homo sapiens

<400> 50
gcatccgcaa gacggtca 18

<210> 51
<211> 20
<212> DNA

<213> Homo sapiens

<400> 51
ttgattatgg cgcccaacag 20

<210> 52
<211> 20
<212> DNA
<213> Homo sapiens

<400> 52
gctggtccta ccccttgtgc 20

<210> 53
<211> 20
<212> DNA
<213> Homo sapiens

<400> 53
cgtgagcgga gcagagagaa 20